A Person- and Time-Varying Vector Autoregressive Model to Capture Interactive Infant-Mother Head Movement Dynamics

Meng Chen, Sy-Miin Chow, Zakia Hammal, Daniel S. Messinger & Jeffrey F. Cohn

View supplementary material

Published online: 12 Jun 2020.

Submit your article to this journal

View related articles

View Crossmark data

Routledge
Taylor & Francis Group

Check for updates

A Person- and Time-Varying Vector Autoregressive Model to Capture Interactive Infant-Mother Head Movement Dynamics

Meng Chen[a], Sy-Miin Chow[a], Zakia Hammal[b], Daniel S. Messinger[c], and Jeffrey F. Cohn[b,d]

[a]Human Development and Family Studies, The Pennsylvania State University; [b]The Robotics Institute, Carnegie Mellon University; [c]Departments of Psychology, Pediatrics, Music Engineering, Electrical and Computer Engineering, University of Miami; [d]Departments of Psychology, Psychiatry, and Intelligent Systems, University of Pittsburgh

ABSTRACT

Head movement is an important but often overlooked component of emotion and social interaction. Examination of regularity and differences in head movements of infant-mother dyads over time and across dyads can shed light on whether and how mothers and infants alter their dynamics over the course of an interaction to adapt to each others. One way to study these emergent differences in dynamics is to allow parameters that govern the patterns of interactions to change over time, and according to person- and dyad-specific characteristics. Using two estimation approaches to implement variations of a vector-autoregressive model with time-varying coefficients, we investigated the dynamics of automatically-tracked head movements in mothers and infants during the Face-Face/Still-Face Procedure (SFP) with 24 infant-mother dyads. The first approach requires specification of a confirmatory model for the time-varying parameters as part of a state-space model, whereas the second approach handles the time-varying parameters in a semi-parametric ("mostly" model-free) fashion within a generalized additive modeling framework. Results suggested that infant-mother head movement dynamics varied in time both within and across episodes of the SFP, and varied based on infants' subsequently-assessed attachment security. Code for implementing the time-varying vector-autoregressive model using two R packages, *dynr* and *mgcv*, is provided.

KEYWORDS
Time-varying parameters; vector autoregressive models; state-space models; generalized additive models; parent-infant interactions; head movements; still-face paradigm

Self-organization is a process through which orderliness emerges from apparent disorder (Lewis & Ferrari, 2001). The idea of self-organization is deeply entrenched in psychology (Bosma & Kunnen, 2011; Kelso, 1995; Magnusson & Cairns, 1996). In the area of human movement, for example, a simple motion involves approximately 10^2 muscle, 10^3 joints, and 10^{14} cells. Yet, as our bodies are capable of self-organization, human movements can be effectively captured by changes in a few key dimensions (Bertenthal, 2007; Turvey, 1990). Newell (1990) proposed a time scale of human action that organizes study interests in the field of psychology into a set of hierarchical levels of analysis. According to Newell's classification scheme, changes occurring in the human body include neural activities that unfold over milliseconds, simple cognitive operations (such as directing attention) that evolve over seconds, rational decision-making processes that unfold over minutes or hours, as well as social processes (e.g. forming a relationship) that

occur on the scales of days, weeks, or even months (Bertenthal, 2007). Self-organization comes into play naturally in change processes that emerge as an integrated result of activities across different levels, and alternatively, time scales.

Interpersonal coordination of movements, including head movements — the focus of our motivating empirical illustration, are characterized by self-organizing change processes that unfold over multiple time scales (Kelso, 1995). Imagine a hypothetical scenario in which two individuals are conversing with each other. In this case, each individual is constantly attending to the other individual's bodily cues such as head nodding/shaking, postures, and other quick second-by-second movements; making periodic inferences of the other individual's emotions or interest level based on these bodily cues; and adjusting his/her own emotions and behaviors accordingly. We may notice that the conversation progresses from delivery of cordial updates to engagement in an exciting — or

even heated — debate, and eventually cools down as the conversation comes to an end. The conversation is thus a self-organizing process that encompasses multiple types of social and cognitive sub-processes at a micro level. The utilization of dynamic systems theory emphasizes both the observed expression of the process as well as its temporal evolution as a whole (Nowak & Lewenstein, 1994). A dynamic systems model thus allows us to extract and formulate key patterns of change as simplified mathematical equations with a manageable set of parameters, that can further be evaluated against empirical data using targeted analytic tools and techniques of choice (van Geert, 2018). The same dynamic system model with different sets of parameters can manifest very different observed trajectories through time. Therefore, allowing the parameters in a model to comprise multi-timescale changes provides one viable way of representing the over-time progression of such a self-organizing process.

Substantial work exists in the econometric, statistical, engineering, as well as social and behavioral sciences literature on longitudinal models with time-varying parameters (TVPs). Varying coefficent models, which were popularized by Hastie and Tibshirani (1993), originally involve cross-sectional models that posit varying relationships between predictors and the outcome as functions of covariates. These varying coefficients are often approximated using spline or functional data analysis methods nonparametrically (completely model-free), or semiparametrically (partially model-free, with spline methods embedded within a model that includes other parametric components). Extensions to the longitudinal context with time as a covariate have gained popularity in the past decade (e.g., Cao et al., 2012; Liang et al., 2010; Wu & Tian, 2018). Increased applications have also emerged in the psychological literature (Bringmann et al., 2017; McKeown & Sneddon, 2014), sometimes under the alternative name of time-varying effect model (TVEM; Li et al., 2014), particularly in examining substance use and intervention-related issues (e.g. Shiyko, Naab, Shiffman, & Li, 2014; Vasilenko et al., 2014). Beyond the spline and functional data literature, variants of discrete-time (e.g., autoregressive models with TVPs; Chow et al., 2010; Del Negro & Otrok, 2008; Harvey, 2001; Molenaar, 1987, 1994; Molenaar et al., 2009; Prado et al., 2001; Rajan & Rayner, 1996; Tarvainen et al., 2006; Wang et al., 2014; Weiss, 1985) and continuous-time models (Chen et al., 2018) with TVPs have also been proposed and estimated within a time series and state-space context, and used to represent a broad range of phenomena from individuals' physiological responses (Molenaar, 1994; Tarvainen et al., 2006), glucose level (Wang et al., 2014), affect (Chow et al., 2009; 2011), and dyadic coupling between individuals (Chow et al., 2010; Molenaar et al., 2009).

In this paper, we consider and illustrate two approaches to implement variations of a vector autoregressive (VAR) model with TVPs, which in the remaining of this paper will be referred to as a time-varying VAR (TV-VAR) model, to evaluate the dynamics of head movements in mothers and infants during the Still Face paradigm (SFP). The first approach, referred to herein as the *state-space modeling* approach, requires specification of a model for the TVPs as part of the dynamical systems model describing the endogenous processes of interest (Chow et al., 2011; Molenaar et al., 1992) — in this case, infant and mother head movements. The TVP model can, in turn, include time-, person-, and/or dyad-specific predictors. Depending on the model adopted for the TVPs, the state-space approach may vary from semi-parametric to strictly parametric (confirmatory) in nature. Here, we specify a theory-driven parametric model under the state-space approach to sharpen its contrasts with the second modeling approach. The second approach handles the TVPs in a semi-parametric (partially model-free) fashion within a *generalized additive modeling* (GAM) framework (Bringmann et al., 2018). Using data from the SFP and two R (R Core Team, 2018) packages, *Dynamic Modeling in R* (dynr; Ou et al., 2019) and *mgcv* (Wood, 2019), we demonstrate and compare the results from the two modeling approaches as well as their respective strengths and weaknesses.

The remainder of this paper is organized as follows. We first present the motivating example to highlight some of the key questions of interest in utilizing head movement to study early parent-infant interactions. We then introduce the TV-VAR model, followed by a description of the adaptations made by us to capture targeted parent-infant interaction changes during the SFP. Next, we present the two approaches for fitting TV-VAR models investigated in the present paper, followed by the corresponding estimation details (with demonstrative R code in Appendix). Then, we present empirical modeling results from the motivating example, and demonstrate how the two modeling approaches can provide distinct but complementary insights on differences in interactive dynamics over time and between dyads. We conclude with some remarks on the contributions and limitations of the two modeling approaches and the empirical study.

Motivating example: head movement dynamics in infant-parent interactions

Parent-child co-regulation is an important aspect of early self-regulation often regarded as a precursor of self-initiated regulatory behaviors in later childhood. In the first three years of life, children progressively gain a variety of abilities that allow them to manage different levels of regulation, from being soothed by parents and other caregivers to self-initiated soothing and control, to active regulation of one's own emotions and behaviors (Kopp, 1982; Rothbart et al., 1992). Before six months of age, self-regulation in infants mainly consists of modulating states of arousal (Calkins, 2011; Feldman, 2003; Kopp, 1982). Before infants can achieve self-initiated regulation, they rely heavily on input and feedback from their parents. At the same time, parents' emotional and behavioral states are affected by their infants' behaviors, valence, and arousal levels (Chow et al., 2010; Cohn & Tronick, 1988; Jaffe et al., 2001). That is to say, in the interactions and co-regulatory processes between infants and parents, the future states of the individual(s) depend on current action and reaction of the individual as well as those of the partner. This is in line with the view of infant-caregiver interaction as a dynamical system, and it is possible to extract and express the complex patterns of synchrony and mismatches in infant-caregiver interaction over time as specific patterns of change over time (van Geert, 2018).

Head movement, just as vocalization and facial expressions, is an important behavioral aspect of emotion communication and social interaction. Previous research has provided evidence that overall movements, including head movements, provide information about the intensity of arousal (Ekman & Friesen, 1974; Kleinsmith & Bianchi-Berthouze, 2013; Wallbott, 1998). Humans often use head movements to convey and detect emotional intensity and meaning (Hammal et al., 2014; 2015; Michel et al., 1992), and they serve special functions in communication practices such as turn-taking and back-channeling (Duncan, 1972; Jokinen et al., 2010; Michel et al., 1992). However, unlike vocalization and facial expressions, which are commonly investigated in research through voice intonation analysis and face recognition tools, information and affect transmission via head movements are often overlooked despite their central roles in human communication.

Advances in automated, unobtrusive, continuous annotation of behavior now make it feasible to gather intensive head tracking data through automated software programs (e.g. Cox et al., 2013; Jeni et al., 2017). Unlike emotional coding schemes that require human coders, automated measures such as head movements are less prone to subjective human biases, and have been shown to be a valid and meaningful alternative to human coding. Research has been sparse in investigating head movements in the context of emotion communications, and especially so in studies of co-regulation. Previous work from Hammal et al. (2015) suggested that quantitative measures of head movements in parent-infant interactions were strongly associated with age-appropriate emotion challenges, thus opening up the new possibilities of using automated head movement measures to uncover characteristics of dyadic dynamics during these interactions.

The current article presents analysis of the data from a previously published study in Hammal et al. (2015) for investigating face-to-face interactions through the channel of head movements under the experimental manipulation of the SFP. The original sample consisted of 42 parent-infant dyads, and 10 more dyads became available for analysis since then, resulting in a total of 52 dyads. The SFP (Tronick et al., 1978) consists of three equal-length (lasting two minutes each) but distinct episodes (Face-to-Face (FF), Still Face (SF) and Reunion (RE)) of parent-infant interaction. It is intended to assess parent-infant reciprocity and infant response to, and recovery from, disturbance of normal dyadic interactive behavior. As briefly as each episode lasts within the SFP, a previous study by Chow et al. (2010) reported substantial over-time variations (non-constancy) in the dynamics between mothers and infants even within the FF and RE episodes based on human rater data. Yet to be clarified, however, are whether such over-time variations are also evidenced in infant-mother interactive head movement dynamics, and the practical implications of such within-episode variations. Thus, the present study seeks to address: (1) within- and between-episode variability in infant-mother interactive head movement dynamics; (2) whether these sources of variability relate to meaningful between-dyad differences, such as attachment outcome; and within-dyad contextual differences, such as under positively as compared to negatively valenced interactions; and (3) consistency in the modeling results as deduced from the GAM vs. the state-space modeling approach.

Data descriptions and preprocessing

To quantify head movement dynamics, a person-independent 3D face tracker (Zface[1]), was used to track the 3 degrees of rigid head movements (i.e., pitch, yaw, and roll) and 49 facial landmarks, or fiducial points, from video recordings on the interactions (Jeni et al., 2017). Head angles in the horizontal

---

[1]The current version of the software is now publicly available at https://github.com/department-of-psychology/AFARtoolbox.

(i.e., pitch), vertical (i.e., yaw), and lateral (i.e., roll) directions were used for analysis. A total of 15% of the video frames could not be tracked. Several conditions contributed to tracking failure, including self-occlusion (hands on the face), extreme head movement, and location change (i.e., child moved out of the frame). Proportions of successfully tracked frames were used for analyses. The raw data contained over 3700 measurement occasions per participant per episode, with a sampling rate of one measurement every 33.366 milliseconds. Our interest is in studying the individuals' conscious and unconscious self-regulation and interactions, which in Newell's time scales of human action, would falls within the "deliberate act" band which is on the unit of 100 milliseconds, or the "operations" band which is on the unit of seconds (Newell, 1990). We were particularly interested in capturing the latter. Thus, we performed data aggregation over every 15 frames to smooth out micro-level noise that may be too fine-grained for the interactive process of our interest, so that the time elapsed between two consecutive measurements was roughly 0.5 second. To ensure sufficient data for ascertaining system dynamics, we included only the dyads meeting the following two criteria for each dyad member: (1) the maximum length of successive missing data points did not exceed 120 (which translates roughly to 60 seconds, namely, at least half of the data from each episode were available); and (2) there existed at least 40 successive observed data points in all of FF, SF and RE episodes after data aggregation. Following these exclusion criteria, 24 dyads were retained from the sample. The average number of non-missing aggregated measurements was 665.2 during the entire SFP, with a minimum of 583 and a maximum of 737 (FF: mean 228.8, min. 153, max. 248; SF: mean 229.5, min. 152, max. 248; RE: mean 221.9, min. 173, max. 248). The mean infant age in this sample was 3.98 months, with a standard deviation of 0.34 months.

For each participant (mother or infant), the three head angle measures (i.e., pitch, yaw, and roll) were then combined into a single variable by calculating the Mahalanobis distances (MDs) from the participant's baseline angles on a reference occasion, defined as the beginning of each episode (i.e., the first available measurement in FF, SF or RE), as:

$$\mathrm{MD}(\boldsymbol{x}_{i,p,t}|ep) = \sqrt{(\boldsymbol{x}_{i,p,t} - \boldsymbol{\mu}_{ep,i,p,1})^\top \boldsymbol{S}_p^{-1}(\boldsymbol{x}_{i,p,t} - \boldsymbol{\mu}_{ep,i,p,1})} \tag{1}$$

where $\boldsymbol{x}_{i,p,t}$ is a vector of observed three dimensions ($[pitch_{i,p,t}, yaw_{i,p,t}, roll_{i,p,t}]$). The i, p, and t subscripts denote, respectively, dyad i, group p (which can

further take on the value of b for infant or m for mother), and time t. The mean or "center point" for distance calculation, $\boldsymbol{\mu}_{ep,i,p,1}$, is a vector containing the pitch, yaw, and roll measures for mother/infant i on the first occasion of each episode for each participant within dyads ($[pitch_{i,p,1}, yaw_{i,p,1}, roll_{i,p,1}]_{ep}$). These initial head angles corresponded to the first instance of valid, front-facing video images of the dyads before other episode-related changes unfolded, and thus served as a practical reference point for our modeling purposes. $\boldsymbol{S}_p$ is the group- (mothers or infants) specific covariance matrix. For infants, we used the covariance of these three measures across all episodes ($\boldsymbol{S}_b = \mathrm{cov}((\boldsymbol{pitch}_b, \boldsymbol{yaw}_b, \boldsymbol{roll}_b)^\top)$). For mothers, we used the covariance matrix of these three measures in the FF session only, $\boldsymbol{S}_m = \mathrm{cov}((\boldsymbol{pitch}_{m,FF}, \boldsymbol{yaw}_{m,FF}, \boldsymbol{roll}_{m,FF})^\top)$, to have a covariance matrix that better reflected mothers' typical ranges of head movements. Person-specific linear trends were then removed from all participants' aggregated head movement within each episode, and the resulting data were re-standardized using the group-wise standard deviations (i.e., with all mothers in one group and all infants in the other) computed using data across all three episodes. We chose to standardize the data across all episodes using the group standard deviations, as opposed to standardizing within each episode with each individual's own within-episode standard deviation, to perserve some between-individual differences in data variability to be captured with the TVPs. In the remaining of the article, the detrending step and scaling step together are referred to as "data preparation" for short.

For illustration purposes, Figure 1 contains the plotted observed head movement data for three randomly sampled dyads through SPF before and after data preparation. The plots indicate that the data preprocessing procedures helped remove some of the arbitrary shifts in head positions as the participants transitioned through the SFP episodes, while also preserving some of the between-dyad and between-episode differences of interest in this study. Furthermore, Figure 2 offers a visual mapping of the resulting MD measures to the original video clip during SF. At the beginning of the SF episode, this specific child was showing minimal movements, which, in turn, resulted in MD values that were around 0. Toward the later half of the episode, the child began to show a greater range of head movements (e.g. lifting head and looking up, turning to the right, edging and looking left in the last three screen shots). These increased head movements were, in turn, evident from the larger absolute MD values.
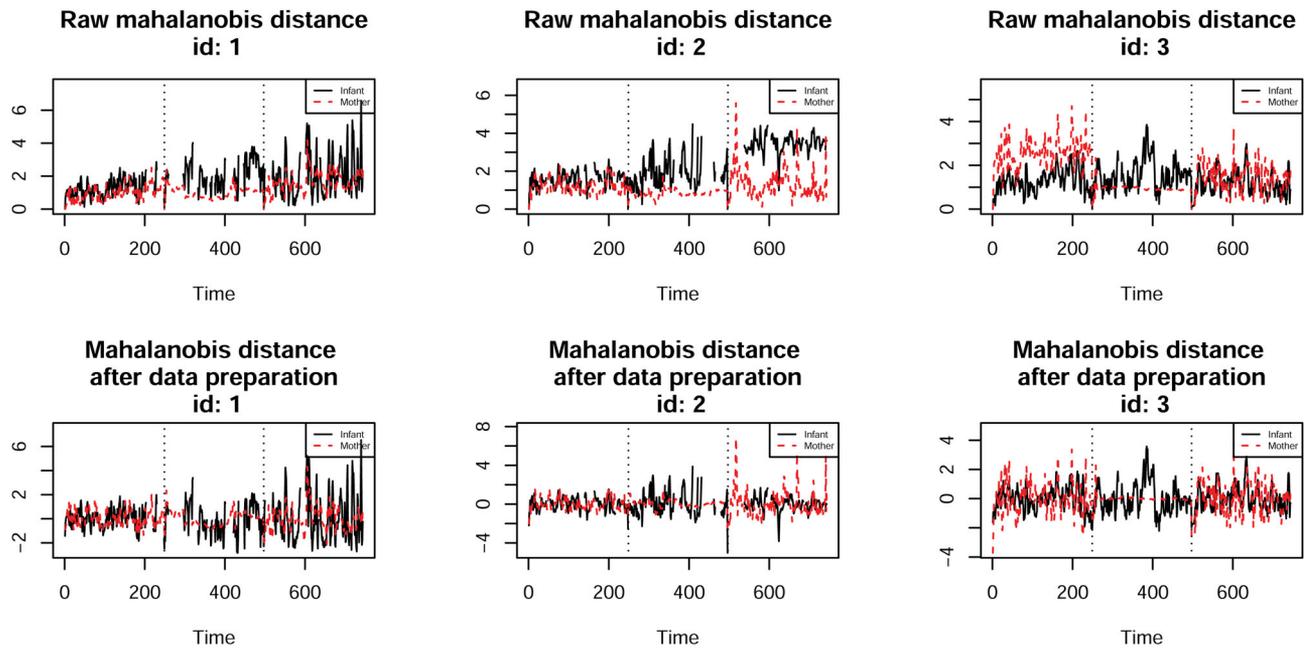
Figure 1. Head movement data before and after data preparation procedures for three randomly selected three dyads.
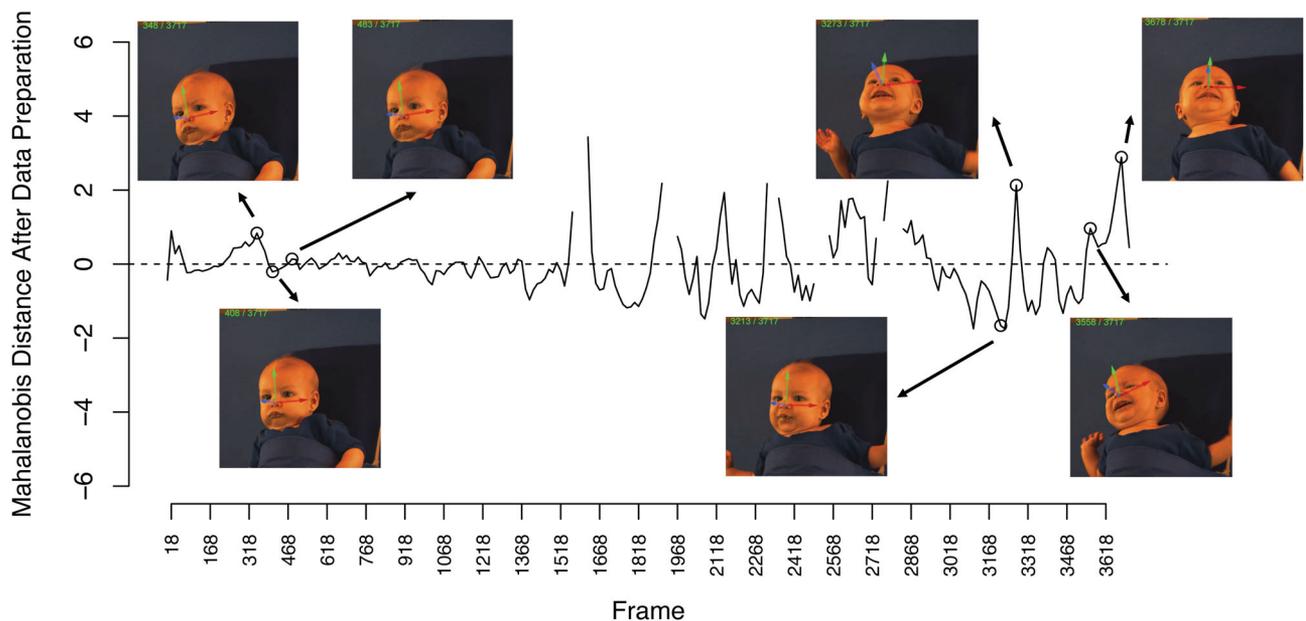
Figure 2. A plot of the MD time series computed for one infant during the SF episode with screen shots from the original video clip. The increases in head movement magnitude during the later half of the episode correspond well with the increases in affect arousal manifested by the infant.

Smiling in parent-infant face-to-face interaction is often studied in the context of emotional communication. Infants tend to respond to mothers' smiling expressions with their own smiles, and seeing the smile for their own mothers elicits a response in the brain region associated with positive affect information and reward mechanism (Minagawa-Kawai et al., 2009). To investigate contextual differences in head movement

dynamics under positively vs. negatively valenced interactions, we used a binary marker of mother smiling by applying a previously validated smile detection classifier (adapted from Girard et al., 2015). Inter-system agreement between the classifier and expert manual annotation was moderate to high ($\kappa = 0.71$).

As a marker of between-dyad differences, we used Richters' Attachment Security Scale (Richters et al.,

1988). This scale provided a continuous index of attachment security derived from expert ratings of the Strange Situation (Ainsworth et al., 1978) administered to the infants in this study at 15 months of age.

TV-VAR models

In this paper, we utilized a series of TV-VAR models to examine our questions of interest. The TV-VAR model is a multivariate discrete-time dynamic system model that allows us to capture patterns of temporal dependencies both within an individual and also between dyadic member simultaneously. In addition, it allows the temporal dependencies to vary over time through incorporation of TVPs. To ease presentation, we begin our illustration with a VAR model with time-invariant parameters.

Imagine a hypothetical scenario in which a mother tries to comfort a frustrated, crying infant. With time, the arousal level of the child is likely to decline if the mother helps the child regulate (e.g. by employing strategies such as distraction or verbal soothing). Despite the mother's help, the child may not calm down instantly, that is, the child would likely exhibit a certain level of continuity of the previous high arousal. The mother's effect, the child's own continuity in arousal, and other sources of stochastic influences from the environment can all be incorporated into a model that describes the fluctuations in the child's arousal levels around a baseline, or a desired level of arousal that reflects the child's own temperament, for instance. A similar process can be applied to the mother in this scenario as well. Observing the infant's crying may increase the arousal level of the mother. The mother's arousal level would also show some continuity in this case as the mother tries to self-regulate. As such, we can also describe the mother's arousal level as a process that fluctuates around her own baseline, and the extent of deviations from baseline, in turn, would exert an influence on the child's deviations in arousal from the child's baseline. Following these characteristics, we can model the arousal levels of the mother and infant over the course of this interaction with the following VAR model of order 1:

$$\begin{bmatrix} infant_{it} \\ mother_{it} \end{bmatrix} = \begin{bmatrix} int_b \\ int_m \end{bmatrix} + \begin{bmatrix} ar_b & cr_{mb} \\ cr_{bm} & ar_m \end{bmatrix} \begin{bmatrix} infant_{i,t-1} - int_b \\ mother_{i,t-1} - int_m \end{bmatrix}$$
$$+ \begin{bmatrix} \zeta_{b,it} \\ \zeta_{m,it} \end{bmatrix},$$
$$\begin{bmatrix} \zeta_{b,it} \\ \zeta_{m,it} \end{bmatrix} \sim N \left(\mathbf{0}, \begin{bmatrix} \psi_b \\ \psi_{bm} & \psi_m \end{bmatrix} \right),$$

(2)

in which i indexes dyad ($i = 1, 2, ..., N$, where N is the total number of dyads in a sample). In the above

scenario, $N = 1$ and t indexes time measured at discrete, equidistant values ($t = 1, 2, ..., T_i$. T_i is the maximum of time index for dyad i). The system variables, $infant_{it}$ and $mother_{it}$, correspond to arousal levels as indicated by head movements in dyad i at time t of the infant and the mother respectively. The intercept parameters, int_m and int_b, describe the levels that the system variables evolve around. The dynamic evolutionary patterns of the system are described by four parameters: ar_b and ar_m, the autoregressive (AR) parameters, along with cr_{mb} and cr_{bm}, the cross-regression (CR) parameters. The components $\zeta_{b,t}$ and $\zeta_{m,t}$, hereby referred to as process noises, represent random disturbances to the system, including disturbances caused by internal or environmental influences that cannot be predicted by knowing the infant and mother previous arousal levels at time t-1.

AR parameters capture the influence of system variables on themselves over time. For example, ar_m describes how much the mother's arousal at the previous observation ($t - 1$) influences the current observation (t). Because the influence of previous observations is limited to that from occasion $t - 1$, Equation 2 depicts a VAR model of order 1, or VAR(1) process. In the affect literature, AR parameters are frequently referred to as inertia (e.g. Kuppens et al., 2010), and in the regulatory literature as self-contingency (e.g. Beebe et al., 2016). It reflects the temporal influence of an individual's state of interest on itself, and thus the continuity of behaviors or emotions. Emotional inertia can be noted as the resistance to change. A high AR value suggests that an individual's current state can be largely predicted using his or her previous state, thus extreme emotions are more likely to persist and less responsive to environmental influence or regulatory efforts. Therefore, emotional inertia is often associated with ineffective emotional regulation and psychological maladjustmnents (Kuppens et al., 2010). Here, we adopt the term inertia to denote the AR parameters to better reflect predictability and rigidity of movements. Further, we use the term *state* throughout broadly to refer to an individual's unobserved underlying process of interest (e.g., emotional valence or arousal).

Figure 3 demonstrates the hypothetical dyadic trajectories of Equations 2 with relatively[2]: (a) high AR values ($ar_m = ar_b = 0.8$), and (b) low AR values ($ar_m = ar_b = 0.5$). In this scenario, a value of 0 represents a dyad member's average level of arousal, higher positive numbers indicate higher arousal than average, and negative numbers represent lower arousal than average. To elucidate the trajectories of the system in

---

[2]R code for simulating data and reproducing Figure 3 is included in the Supplementary Material.
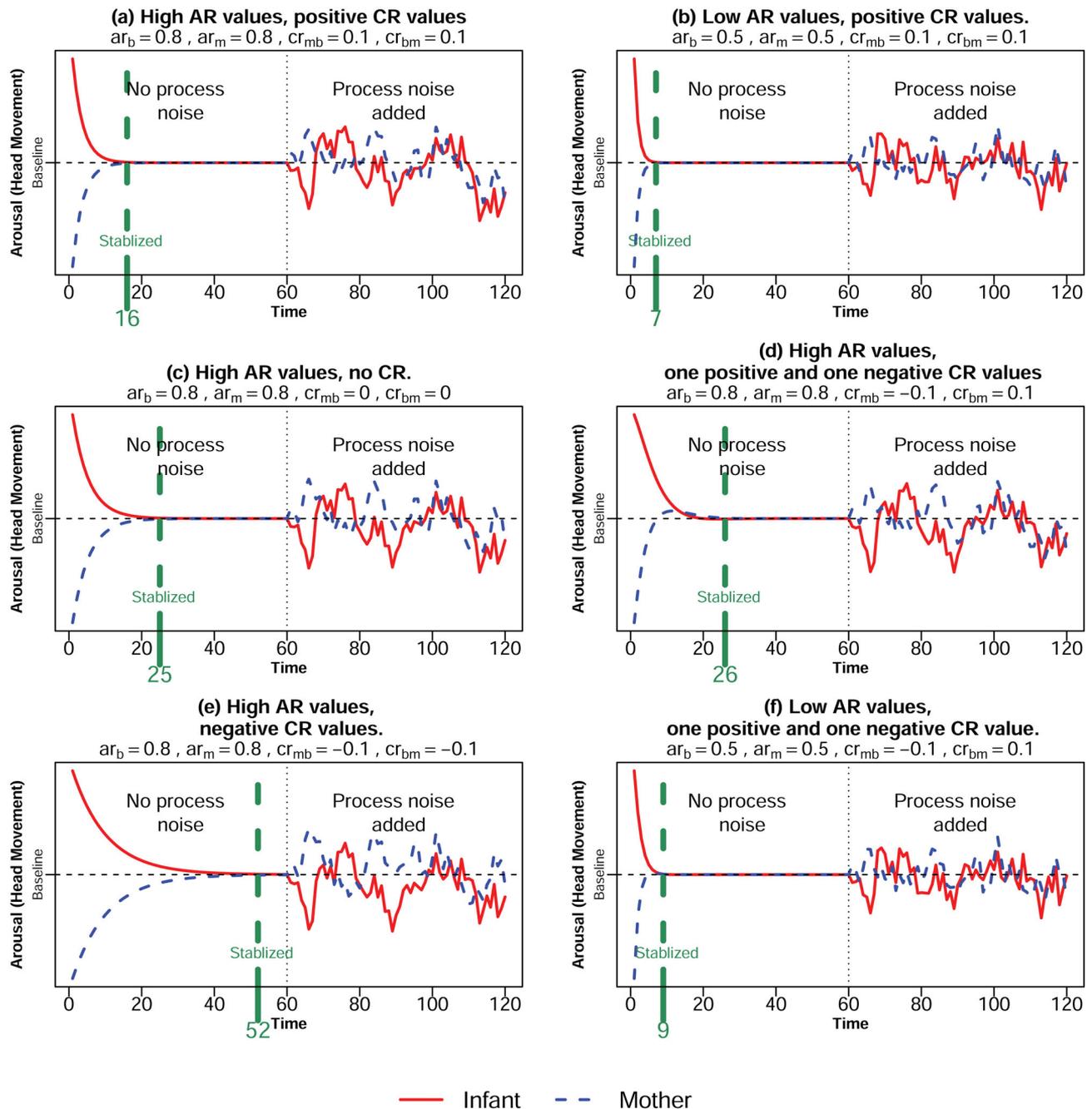
Figure 3. Realizations of the same VAR model in Equation 2 can look very different depending on the parameter values. No process noise is added before the 60th time unit so that the trajectories are entirely driven by the AR and CR parameters. The vertical dashed green line and the associated value on the x-axis represent the time when both dyadic members arrive at a stable state. After the 60th time unit (vertical dotted line), the identical sequence of bivariate process noises are added to trajectories in each plot.

the presence/absence of process noises, we added process noises starting only after time = 60. If the two members start from states far away from their desired stable states (in Figure 3 set to a value of zero), it takes the dyad 16 time units in the high AR scenario, but only 7 time units in the low AR scenario, for the dyad members to return to their stable state in the absence of new process noises. In other words, under low AR values, mother and infants are less resistant to

change, thus making the other person's influence more salient. From the 60th time unit on, an identical two-dimensional sequence of process noises is added to all the systems in Figure 3. Here the trajectories become "rougher" and are harder to predict based on information from time $t - 1$. That is, the inclusion of process noises has now made these processes *stochastic*. The high AR scenario in (a) is characterized by longer and more extreme bouts of ebb and flow

compared to the low AR scenario (b), where the fluctuations are smaller in magnitudes and clutter more closely around the baseline of zero.

The CR parameters capture the influence of system variables on each other. For example, in Equation 2, cr_{bm} indicates how much the infant's arousal at the previous observation affects the mother's arousal at the current observation. The CR parameters can capture the relation in behaviors between two dyad members as indicated by Granger causality (Molenaar, 2019), after taken into consideration the continuity carried by the members themselves (as accounted for by the AR parameters). Beebe et al. (2016) also referred to this type of relation as "interactive contingency" because it reflects one member's adjustment in state relative to the other member's prior state. The subfigures (a), (c), (d), and (e) in Figure 3 represent four scenarios of Equations 2, each with a different set of the CR parameter values: (a) positive for both parameters, (c) zero for both CR parameters (thus no interactive contingency at all), (d) positive value for one CR parameter and negative value for the other, and (e) negative for both parameters. In (c), there is no interactive contingency at all, as in a scenario where the mother and infant show no interactive contingency with respect to each other whatsoever. It takes 25 time units for both trajectories to converge at a state level (zero) without any noise or disturbance to the system. Subfigure (a) represents a scenario where the trajectories are pulled toward each other, yielding a more efficient co-regulated system that converges to their stable states faster (16 time units) compared to when no interactive contingency is present. In contrast, (e) represents a scenario where there is a weak "anti-regulatory" force against each other. In such a scenario, when the mother is trying to respond to the infant's heightened emotional arousal with a lower level of arousal, conditional on the fact that their arousal states started from opposite directions, such a mismatch in arousal levels and the "anti-regulatory force" create further delays (requiring 52 time units) for the two members to calm down toward their stable states. The last scenario, (d), is where there is a repelling force on the infant from the mother, and a pulling force on the mother from the infant. This might mirror the case where the infant is resistant to the soothing actions of the mother, and meanwhile the mother's arousal level is brought up by the infant not cooperating. Without any process noise, this system still converges to the stable level at time unit 26.

When within-dyad differences across time exist, for example, when the interaction dynamics changed when entering SF, the system then violates stationarity, which is a key assumption of time-series analysis. Most commonly used definition of stationarity, the covariance stationarity, states that the first and second moments of the time-series data should be time-invariant. Translated into terms in Equation 2, this implies that the intercepts and the AR, CR dynamic parameters need to be time-invariant. One way to account for within-dyad variations over time, and sometimes also between-dyad differences, in dynamics using the VAR model is to allow for TVPs. In other words, we can have the four dynamic parameters (i.e. AR and CR parameters) to differ across dyads and over time. Here we use the ar_b parameter in Equation 2 as an example. We can replace it with a time-and dyad-specific version,

$$ar_{b,it} = f(t, ar_{b,i,t-1}, \mathbf{v}_i, \mathbf{x}_{it}) + \zeta_{ar_b,i,t}$$
$$\zeta_{ar_b,i,t} \sim \mathbf{N}(0, \psi_{ar_b}),$$
(3)

where $ar_{b,it}$ is a function of the following components: time (t), the value of the AR parameter at time $t-1$, $ar_{b,i,t-1}$, a vector of dyad-specific characteristics ($\mathbf{v}_i$), a vector of time-specific predictor variables ($\mathbf{x}_{it}$). $\mathbf{v}_i$ may include characteristics of individual members that constitute dyad i – in our case, the attachment level of the infant as indicated by the Richter scores. $\mathbf{x}_{it}$ in our motivating example is a one-dimensional exogenous time-varying binary covariate, mother's smile. f can be parametric or nonparametric. The term $\zeta_{ar_b,i,t}$, which is usually assumed to conform to a normal distribution, represents residual or process noise that account for deviations from the predicted AR. If f contains only an intercept parameter ($f(.) = int_{arb}$) and no process noise is added, then $ar_{b,it}$ in Equation 3 is equivalent to the time-invariant ar_b parameter in Equations 2. In fact, in model fitting involving any potential TVP, a model with only process noise is often fitted to that specific parameter before any theory-driven models to see whether representing such parameter as a TVP is necessary. If the estimated process noise variance is different from zero, then it provides evidence that suggests there is enough variability in the parameter and thus may be an indication that the parameter varies through time. Otherwise, the parameter would be specified as time-invariant (Chow et al., 2011).

Although in the previous paragraph we used an AR parameter as an example, AR parameters are not the only ones that can be time-varying. In fact, previous work by Chow et al. (2010) found time variations in the concurrent association between infants and parents during FF and RE episodes using a stochastic regression

model predicting infant emotional valence while controlling for previous infant valence. These results provide initial support for incorporating CR parameters as TVPs but do not treat the dyad as a bivariate system, nor do they model potential TVP covariates.

In summary, the standard VAR model defines (successful) regulation and thus co-regulation of the pair as dependent not only on the affect of the other member (CR parameters), but also on the resistance (or reversely, susceptibility) to change of the individual him/herself (AR parameters). A thorough investigation of co-regulation requires consideration of both interactive contingency and inertia. Therefore, to study the phenomenon of co-regulation through the VAR model, it is important to take the entire evolutionary pattern of the system into consideration, as both AR and CR parameters provide unique but complementary pieces of information concerning the dynamics of dyadic interactions. TV-VARs extend on the traditional VAR model with possibilities to account for any between-dyad and across-time differences in the dynamic system patterns. One novel contribution of the article is to illustrate a rather systematic investigation of the dynamic parameters in a VAR model that may be time-varying in the context of intensive longitudinal data on parent-infant interactions, as well as the associations between variations of dynamics (marked by TVPs) and dyad-specific characteristic (infants' later attachment) and time-varying environmental factor (mothers' smiles).

Estimation details for TV-VAR models

State-space approach

The state-space modeling approach operates by incorporating the TVPs as additional latent variables in the context of a state-space model, and subsequently estimating the over-time fluctuations in the TVPs with other latent variables in the system. The specific form of state-space model we consider in this study comprises a dynamic model expressed as:

$$\boldsymbol{\eta}_{it} = \boldsymbol{f}_{\boldsymbol{\eta}}(\boldsymbol{\eta}_{i,t-1}, \boldsymbol{\beta}) + \boldsymbol{\zeta}_{it}, \quad \boldsymbol{\zeta}_{it} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{\zeta}), \quad (4)$$

and a linear measurement function written as:

$$\boldsymbol{y}_{it} = \boldsymbol{\Lambda}\boldsymbol{\eta}_{it} + \boldsymbol{\epsilon}_{it}, \quad \boldsymbol{\epsilon}_{it} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{\epsilon}). \quad (5)$$

In the above model formulation, $\boldsymbol{\eta}_{it}$ is a p-dimensional vector representing latent variables — also referred to as "states" in the state-space literature — for the system unit i at time t; $\boldsymbol{f}_{\boldsymbol{\eta}}$ ($\mathbb{R}^p \to \mathbb{R}^p$) is the state transition function from a previously time $t-1$ to the current time t; and $\boldsymbol{\beta}$ is a k-dimensional vector of parameters

in $\boldsymbol{f}_{\boldsymbol{\eta}}$. $\boldsymbol{\zeta}_{it}$ is also a p-dimensional vector, and it represents the process noise at time t. $\boldsymbol{y}_{it}$ is a q-dimensional vector of observed manifestations of the latent states $\boldsymbol{\eta}_{it}$; $\boldsymbol{\Lambda}$ is a $q \times p$ matrix of the measurement loading that links the latent states $\boldsymbol{\eta}_{it}$ to the observed $\boldsymbol{y}_{it}$; and $\boldsymbol{\epsilon}_{it}$ is a q-dimensional vector representing measurement errors. In the traditional VAR model, the latent states are the system variables of interest $\left(\begin{bmatrix} infant_{it} \\ mother_{it} \end{bmatrix} \right.$ in Equation 2 for this study; $p = 2 \Big)$. In the case of TV-VAR, for example, if we are to include a time-varying version of ar_b as illustrated in Equation 3, we add another dimension in $\boldsymbol{\eta}_{it}$ that now it becomes $\begin{bmatrix} infant_{it} \\ mother_{it} \\ ar_{b,it} \end{bmatrix}$ with $p = 3$ and include a parametric model for $ar_{b,it}$ in $\boldsymbol{f}_{\boldsymbol{\eta}}$. $\boldsymbol{\zeta}_{it}$ then is also three-dimensional and becomes $\begin{bmatrix} \zeta_{b,it} \\ \zeta_{m,it} \\ \zeta_{ar_b,it} \end{bmatrix}$, and the corresponding covariance matrix $\boldsymbol{\Sigma}_{\zeta}$ is $\begin{bmatrix} \psi_b & & \\ \psi_{bm} & \psi_m & \\ 0 & 0 & \psi_{ar_b} \end{bmatrix}$. In our particular example, the observed $\boldsymbol{y}_{it}$ is also the two-dimensional latent state variable $\begin{bmatrix} infant_{it} \\ mother_{it} \end{bmatrix}$ representing mother's and infant's head movement. Therefore, $q = 2$ and with the inclusion of $ar_{b,it}$, $\boldsymbol{\Lambda} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, with no measurement error $\boldsymbol{\epsilon}_{it}$ involved.

We used the R package *dynr* (Ou et al., 2018), to implement this approach. In *dynr*, estimation of both the latent states in TV-VAR and time-invariant parameters (sometimes referred to as "dual-estimation") calls for three steps, which we describe briefly in turn below. As an overview, these steps include a filtering step to estimate the values of the latent variables (including the TVPs) and the uncertainty associated with those estimates at time t. Filtering assumes that the observed data are only available up to time t, and the unknown parameters are fixed at their specified (e.g., starting) values. The filtering then leads to by-products that can be used to compute a raw data likelihood function. Optimization of this raw data likelihood function with respect to the unknown parameters essentially involves repeated execution of the filtering step at different parameter values until some pre-defined convergence criteria are met. At convergence, the final (converged) parameter estimates are used to run filtering one more time, followed by a smoothing procedure to generate refined

estimates of the latent variables scores using data from all the time points.

Step 1. Filtering. To perform filtering on the latent variables in Equations (4)-(5), the extended Kalman filter (EKF), an extension of the linear Kalman filter (Kalman, 1960) for nonlinear, discrete-time dynamic models was used. As in the Kalman filter procedure, the EKF involves iterations of prediction steps and update steps over all subjects and time points to yield filtered estimates of the state variables $\boldsymbol{\eta}$ and the associated variance covariance structure $\boldsymbol{P}$ (Anderson & Moore, 1979). For the filtering step, the collection of time-invariant parameters $\boldsymbol{\theta}$ (consisting of elements in $\{\boldsymbol{\beta}, \boldsymbol{\Sigma}_\zeta, \boldsymbol{\Sigma}_\epsilon, \boldsymbol{\Lambda}\}$) are assumed known. Estimation of these parameters are addressed in Step 2 with the optimization algorithm.

Let $\hat{\boldsymbol{\eta}}_{i,t|t-1}$ denote the estimated $\boldsymbol{\eta}_{it}$ at the prediction step, which utilizes information in the data up to time $t-1$, and $\hat{\boldsymbol{\eta}}_{i,t-1|t-1}$ denote the estimated $\boldsymbol{\eta}_{i,t-1}$ at the previous update step, which also utilizes information up to time $t-1$. $\boldsymbol{P}_{i,t|t-1}$ and $\boldsymbol{P}_{i,t-1|t-1}$ represent, respectively, their associated covariance matrices. In each iteration of EKF, the prediction step builds up on results from the previous iteration's update step, which are derived from observations $\{\boldsymbol{y}_1, ..., \boldsymbol{y}_{t-1}\}$, to yield:

$$\hat{\boldsymbol{\eta}}_{i,t|t-1} \overset{\Delta}{=} E(\boldsymbol{\eta}_{i,t}|\boldsymbol{y}_{i,1}, ..., \boldsymbol{y}_{i,t-1}) = \boldsymbol{f}_\eta(\hat{\boldsymbol{\eta}}_{i,t-1|t-1}) \quad (6)$$

$$\begin{aligned} \boldsymbol{P}_{i,t|t-1} &\overset{\Delta}{=} cov(\boldsymbol{\eta}_{i,t}|\boldsymbol{y}_{i,1}, ..., \boldsymbol{y}_{i,t-1}) \\ &= \boldsymbol{J}_f(\hat{\boldsymbol{\eta}}_{i,t-1|t-1})\hat{\boldsymbol{P}}_{i,t-1|t-1}\boldsymbol{J}_f(\hat{\boldsymbol{\eta}}_{i,t-1|t-1})^\top + \boldsymbol{\Sigma}_\zeta, \quad (7) \end{aligned}$$

in which $\boldsymbol{J}_f(\hat{\boldsymbol{\eta}}_{i,t-1|t-1})$ is the Jacobian matrix of $\boldsymbol{f}_\eta$, with element in the jth row and kth column being the first-order partial derivative of the jth function in $\boldsymbol{f}_\eta$ with respect to the kth variable in $\boldsymbol{\eta}_{it}$, evaluated at the most current estimate $\hat{\boldsymbol{\eta}}_{i,t-1|t-1}$.

The prediction step estimates $\hat{\boldsymbol{\eta}}_{i,t|t-1}$ and $\boldsymbol{P}_{i,t|t-1}$ are then carried into the update step as information of $\boldsymbol{y}_{it}$ is utilized to further refine the state estimates as:

$$\begin{aligned} \boldsymbol{v}_{it} &\overset{\Delta}{=} \boldsymbol{y}_{it} - E(\boldsymbol{y}_{it}|\hat{\boldsymbol{\eta}}_{i,t|t-1}) = \boldsymbol{y}_{it} - \boldsymbol{\Lambda}\hat{\boldsymbol{\eta}}_{i,t|t-1} \\ \boldsymbol{V}_{it} &\overset{\Delta}{=} cov(\boldsymbol{v}_{it}) = \boldsymbol{\Lambda}\boldsymbol{P}_{i,t|t-1}\boldsymbol{\Lambda}^T + \boldsymbol{\Sigma}_\epsilon \\ \boldsymbol{K}_{it} &= \boldsymbol{P}_{i,t|t-1}\boldsymbol{\Lambda}^T\boldsymbol{V}_{it}^{-1} \\ \hat{\boldsymbol{\eta}}_{i,t|t} &= \hat{\boldsymbol{\eta}}_{i,t|t-1} + \boldsymbol{K}_{it}\boldsymbol{v}_{it} \\ \boldsymbol{P}_{i,t|t} &= \boldsymbol{P}_{i,t|t-1} - \boldsymbol{K}_{it}\boldsymbol{\Lambda}\boldsymbol{P}_{i,t|t-1} \end{aligned}$$

Here $\boldsymbol{v}_{it}$ is referred to as prediction error at time t, and $\boldsymbol{V}_{it}$ is its variance. Both $\boldsymbol{v}_{it}$ and $\boldsymbol{V}_{it}$ are used in the calculation of likelihood for parameter optimization in Step 2. $\boldsymbol{K}_{it}$, called the Kalman gain, can be seen as a relative weight between variability of the

predicted state estimates, the magnitudes of which are captured by $\boldsymbol{P}_{i,t|t-1}$, and the total variability of the new observations, the magnitudes of which depend both on the variability of the predicted state estimates, $\boldsymbol{P}_{i,t|t-1}$, and also the measurement error covariance matrix, $\boldsymbol{\Sigma}_\epsilon$. The noisier the observations are, the lower the Kalman gain value is. Thus less weight is given to the new observations when updating the state predicted estimates. The current estimate of $\boldsymbol{\eta}_{it}$ from the update step and its associated covariance, $\hat{\boldsymbol{\eta}}_{i,t|t}$ and $\boldsymbol{P}_{i,t|t}$ are subsequently used in the next iteration of prediction step for $\boldsymbol{\eta}_{i,t+1}$. $\boldsymbol{P}_{i,t|t}$ helps quantify the "errors" or uncertainty in the state estimates after new data are available from time t, and is sometimes referred to as the conditional state error covariance matrix (Anderson & Moore, 1979).

Step 2. Parameter Estimation. Parameter estimation is performed by finding parameter estimates that maximize a raw log-likelihood function, also known as the prediction error decomposition function, that can be computed using by-products from the filtering step. The raw data log-likelihood function is expressed as (Chow et al., 2007; Schweppe, 1965):

$$\log l(\boldsymbol{\theta}) = -\frac{1}{2}\sum_{i=1}^N\sum_{j=1}^{T_i}\left[\log(2\pi) + \log|\boldsymbol{V}_{i,t_j}| + \boldsymbol{v}_{i,t_j}^\top\boldsymbol{V}_{i,t_j}^{-1}\boldsymbol{v}_{i,t_j}\right].$$

$$(8)$$

The optimization algorithm employed by *dynr* is a sequential quadratic programing algorithm (Kraft, 1988, 1994) from the open-source library for nonlinear optimization, NLOPT (Johnson, 2014; Ou et al., 2018).

Step 3. Smoothing. The filtering step only uses information up to time t for estimation of the latent states and covariance structure, $\hat{\boldsymbol{\eta}}_{i,t|t}$ and $\boldsymbol{P}_{i,t|t}$. We can further refine these estimates using information contained in observations from the entire time-series, including those from time $t+1$ and so on, via the fixed interval smoother that is run backward in time (Ansley & Kohn, 1985; Chow et al., 2010; Harvey, 2001):

$$\hat{\boldsymbol{\eta}}_{i,t|T} = \hat{\boldsymbol{\eta}}_{i,t|t} + \tilde{\boldsymbol{P}}_{it}(\hat{\boldsymbol{\eta}}_{i,t+1|T} - \hat{\boldsymbol{\eta}}_{i,t+1|t}), \quad (9)$$

$$\boldsymbol{P}_{i,t|T} = \boldsymbol{P}_{i,t|t} + \tilde{\boldsymbol{P}}_{it}(\boldsymbol{P}_{i,t+1|T} - \boldsymbol{P}_{i,t+1|t})\tilde{\boldsymbol{P}}_{it}, \quad (10)$$

where $\tilde{\boldsymbol{P}}_{it} = \boldsymbol{P}_{i,t|t}\boldsymbol{J}_f(\hat{\boldsymbol{\eta}}_{i,t|t})^\top[\boldsymbol{P}_{i,t+1|t}]^{-1}$. This step yields our final estimates of the latent states, including the TVPs, The square roots of the diagonal elements in $\boldsymbol{P}_{i,t|T}$ are the standard deviations of the state estimates after all the data up to time T have been used for estimation, and they can be used as standard errors to form confidence intervals around $\hat{\boldsymbol{\eta}}_{i,t|T}$. When TVPs

are included as additional latent variables in $\boldsymbol{\eta}_{it}$, the pertinent elements in $\hat{\boldsymbol{\eta}}_{i,t|T}$ serve as estimates of the TVPs, with confidence intervals constructed using the corresponding square root elements in $\boldsymbol{P}_{i,t|T}$.

A unique feature of the fitting TV-VAR in the state-space framework is the capability for researchers to specify particular functional forms for TVPs. In theory-driven TVPs, doing so can help confirm any particular pattern of time-varyingness of the parameters and establish connections between time-varying characteristics of dynamics and other factors that may have influences on the dynamics (e.g. Chow et al., 2009; Tarvainen et al., 2004). Even in cases where theories guiding the nature of TVPs are lacking, one can choose functions that are flexible enough as a first probe for TVPs before making decisions about more targeted confirmatory models (e.g. Asparouhov et al., 2018; Chen et al., 2018) or explicitly incorporate nonparametric functions or splines for TVPs (e.g. Tarvainen et al., 2006; Zhu & Wu, 2007).

GAM framework approach

Another approach adopted in the present article to estimate variations of the TV-VAR model is a GAM framework approach. Following Bringmann et al. (2018), we utilized GAMs through the R package, Mixed GAM Computation Vehicle with Automatic Smoothness Estimation (*mgcv*, Wood, 2019) to estimate the over-time trajectories of the TVPs by means of penalized regression splines. A GAM with dependent variable y and predictors $\boldsymbol{x}$ is generally written as (Yee, 2015):

$$
E(y_i) = g(\mu(\boldsymbol{x}_i))
$$

$$
= \beta_1 + \sum_{j=1}^{J} f_j(x_{ij}) + \sum_{k=J+1}^{K} f_k(x_{ik_1}) x_{ik_2}
$$

$$
+ \sum_{h=K+1}^{H} f_h(x_{ih_1} x_{ih_2}) \tag{11}
$$

where β_1 and f_d $(d = 1, ..., H)$ are smooth functions, which in *mgcv* are based on thin plate regression splines[3] by default. The term $f_j(x_{ij})$ represents the smooth functional effect of the jth predictor x_{ij}, an example being a nonlinear time trend. The term $f_k(x_{ik_1}) x_{ik_2}$ allows the effect of the k_1^{th} predictor, x_{ik_1}, on y to vary as a function of the k_2^{th} predictor (in our case, time). This is the key term utilized in the present

article to allow the effects of the lag-1 predictors, $infant_{i,t-1}$ and $mother_{i,t-1}$, to be time-varying. Finally, $f_h(x_{ih_1} x_{ih_2})$ is a tensor product term that allows for approximations of jointly nonlinear effects involving both x_{ih_1} and x_{ih_2}. Tensor product terms are not used in the present article, but see Chow (2019) for examples of modeling with this term.

Using *mgcv*, we considered an alternative TV-VAR model adapted from Equations 2—3 as:

$$
\begin{bmatrix} infant_{it} \\ mother_{it} \end{bmatrix} = \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix} + \begin{bmatrix} f_3(t) & f_4(t) \\ f_5(t) & f_6(t) \end{bmatrix} \begin{bmatrix} infant_{i,t-1} \\ mother_{i,t-1} \end{bmatrix}
$$

$$
+ \begin{bmatrix} \zeta_{b,it} \\ \zeta_{m,it} \end{bmatrix}, \quad \begin{bmatrix} \zeta_{b,it} \\ \zeta_{m,it} \end{bmatrix} \sim N \left(\boldsymbol{0}, \begin{bmatrix} \psi_b \\ \psi_{bm} & \psi_m \end{bmatrix} \right), \tag{12}
$$

and the two system variables, $infant_{i,t}$ and $mother_{i,t}$, need to be manually manipulated to create the lag-1 predictors, $infant_{i,t-1}$ and $mother_{i,t-1}$, to be entered into the regression as predictors. The smooth functions $f_1(t)$ and $f_2(t)$ correspond to time-varying intercept parameters for mothers and babies, respectively; $f_3(t)$ and $f_6(t)$ correspond to time-varying AR parameters ar_{bt} and ar_{mt}, and finally $f_4(t)$ and $f_5(t)$ correspond to time-varying CR parameters $cr_{mb,t}$ and $cr_{bm,t}$. Compared to the original TV-VAR model shown in Equations 2—3, one notable difference is the inclusion of the time-varying intercept terms, $f_1(t)$ and $f_2(t)$ in the model. In this case, a researcher may opt to capitalize on the nonparametric strengths of the GAM framework to simultaneously model other unspecified time trends in all mothers' and infants' trajectories with relative ease.

Let $\boldsymbol{\beta}_d$, $d = 1, ..., 6$, represent the vector of basis coefficients in the smooth function f_d in Equation 12 and $\boldsymbol{\beta}' = (\boldsymbol{\beta}_1', ..., \boldsymbol{\beta}_6')$. The estimated $\boldsymbol{\beta}$ are then obtained by maximizing the penalized log-likelihood:

$$
\log l(\boldsymbol{\beta}) - \frac{1}{2} \sum_{d=1}^{6} \lambda_d \boldsymbol{\beta}_d' \boldsymbol{S}_d \boldsymbol{\beta}_d, \tag{13}
$$

where λ_d is a penalty parameter that controls the importance of smoothness of the approximation curve, f_d, and $\boldsymbol{S}_d$ is the "wiggliness" penalty matrix that defines the smoothness criterion for the dth approximation curve, f_d, the deviations from which are penalized to ensure the smoothness of f_d. The collection of penalty parameter $\boldsymbol{\lambda} = \{\lambda_1, ..., \lambda_6\}$ needs to be selected with care to maintain a balance between goodness-of-fit (measured by the first term in Equation 13) and wiggliness (measured by the second term in the equation). In *mgcv*, $\boldsymbol{\lambda}$ and basis coefficients $\boldsymbol{\beta}$ are jointly optimized through a procedure with nested iterations. The outer iteration handles

---

[3]In thin plate regression splines, the basis is obtained through eigen-decomposition of a data-determined matrix. Please refer to Wood (2003) for details.

optimization of λ using criteria such as generalized cross-validation (GCV) and restricted maxmimum likelihood (REML). Nested within this outer iteration is the estimation of β by using a Newton algorithm to maximize Equation 13 (Wood, 2019). If the dependent variable follows a multivariate normal distribution, as in the case of our model assumption, an approximated REML is used, which showed better performance than GCV in a previous simulation study (Wood, 2011).

Apart from λ, the smoothness in GAM is also affected by the number of basis functions (k). On choosing the value of k, authors of the *mgcv* package suggested running a gam.check() on a fitted GAM to test whether the number is adequate. The test is based on computing an estimate of the residual variance after ordering the residuals according to the predictor values and taking differences of successive residuals. If the value of this estimate divided by the residual variance falls below 1, the residuals are likely to contain patterns not already accounted for by the existing basis expansion and one may consider doubling the value of k and re-fittig GAM (Wood, 2019). After the parameter estimation, *mgcv* uses a Bayesian approach for deriving standard errors of predictions (and confidence bands; see Marra & Wood, 2012, for details), and also for testing the significance of the smooth terms. Significance in *mgcv* is defined against the null hypothesis that a particular smooth term f_d is zero (Wood, 2013).

The approach of TV-VAR model fitting under the GAM framework does not rely on preexisting speculations on the nature and shape of the TVPs and therefore offers a relatively model-free approach to TVP estimations. However, it makes customized specification of theory-driven TVP models difficult for the same reason. One key advantage of the state-space model approach over the GAM approach is the capability to include a measurement structure (Equation 5), albeit not highlighted in the current study. Another difference between these two model fitting approaches lies in how they handle between-unit (e.g. dyad) differences. The state-space model approach aims to extract a universal pattern within the sample by constraining the time-invariant parameters and the general predefined model for TVPs to be the same across dyads, while preserving some between-dyad differences in dynamics as reflected through the process noise elements and the individual filtering and smoothing procedures. Thus, even with the same general model for TVPs, the estimated TVP trajectories would still differ from one dyad to the next. In contrast, in the GAM framework approach, one needs to fit a group-based model to extract, for all dyads in the sample, a common trend for each TVP; alternatively, one may adopt a dyad-specific approach and fit a model separately to each dyad's data. In the case of group-based model fitting, the implied TVP trajectory would be identical across dyads. Either way, incorporating theory-driven parametric functions that link known covariates to the TVPs is not very straightforward in *mgcv*, by the design and nature of the package. Other spline packages exist and have other unique strengths that are beyond the scope of the specific empirical illustration targeted in this article. We provide a brief synopsis in the Discussion section.

Empirical results

The Empirical Results section is organized as follows: preliminary results are first reported to showcase the similarities and differences in results using the aggregated MDs on head movement compared with the previous published results using pitch, yaw and roll separately; then results from the state-space approach are presented, following a process of screening for TVP to confirmatory model fitting with dyad- and time-specific elements predicting TVPs with special attention given to: a) whether CR and AR parameters are time-varying, and if they are, whether the TVPs of CR can be predicted by episodes and b) whether the TVPs can be predicted by the mother's smiles and Richters' Attachment Scale the infant; and finally we represent results from GAM approach and how the conclusions drawn from the GAM approach may be similar or different from those from the state-space approach.

Preliminary results on mahalanobis distances

Descriptive statistics of the MDs measures before and after detrending with across-episode group-wise standardizations, are shown in Table 1. The aggregated MDs between mothers and infants showed only low to moderate concurrent and lag-1 associations, and high lag-1 autocorrelations (Table 2). The lag-2 partial correlations were diminished in magnitude but significant for 51.4% of both mothers' and infants' time series. Both mothers' and infants' post-preparation data exhibited a certain level of nonstationary by the unit root tests (proportion of nonstationary data: infants: 13.89% by the augmented Dickey-Fuller test, 2.78% by the KPSS test for level stationarity; mothers: 23.61% by the augmented Dickey-Fuller test, 9.72% by the KPSS test for level stationarity).

Table 1. Mahalanobis distance (before and after data preparation) across dyads summarized by person and episode.

Before Data Preparation ("Raw" Mahalanobis Distance)

	Mother		
Episode	Mean (SD)	Median	(Min., Max.)
FF	1.413 (0.860)	1.233	(0.000, 7.967)
SF	1.491 (0.918)	1.260	(0.000, 6.829)
RE	1.852 (1.005)	1.717	(0.000, 6.546)

	Infant		
Episode	Mean (SD)	Median	(Min., Max.)
FF	1.498 (1.124)	1.211	(0.000, 7.824)
SF	1.503 (0.992)	1.272	(0.000, 5.902)
RE	1.576 (0.834)	1.576	(0.000, 6.578)

After Data Preparation

	Mother		
Episode	Mean (SD)	Median	(Min., Max.)
FF	0.000 (1.028)	−0.089	(−4.630, 10.435)
SF	0.000 (0.784)	0.016	(−4.858, 9.410)
RE	0.000 (1.125)	−0.050	(−5.485, 6.856)

	Infant		
Episode	Mean (SD)	Median	(Min., Max.)
FF	0.000 (0.981)	−0.052	(−8.115, 5.958)
SF	0.000 (1.095)	−0.088	(−3.499, 6.563)
RE	0.000 (0.985)	−0.065	(−5.026, 6.540)

Table 2. Average correlations and autocorrelations after data preparation.

	$Mother_t$	$Infant_{t-1}$	$Mother_{t-1}$
All Episodes			
$Infant_t$	0.079	0.732	0.042
$Mother_t$		0.077	0.761
FF			
$Infant_t$	0.121	0.746	0.079
$Mother_t$		0.122	0.695
SF			
$Infant_t$	0.083	0.719	0.051
$Mother_t$		0.065	0.751
RE			
$Infant_t$	0.081	0.709	0.029
$Mother_t$		0.079	0.761

Prior to dynamic model fitting, we conducted preliminary analyses to examine the veracity of the MD measure and the effect of omitting dyads with excessive missingness in the time series data. The goal here was to compare the results reported by Hammal et al. (2015) for a larger sample of 42 dyads in two summary measures: displacement and velocity of head movement. Unlike Hammal et al. (2015), who performed separate analyses of displacement and velocity for pitch, yaw and roll for all 42 dyads, the results reported here were based on the combined measure of MD, and only from a subset of 24 dyads with sufficient data for subsequent dynamic modeling. In addition, the head movement measures were extracted in the current study using a different tracker from the one used in Hammal et al. (2015).

Between-episode differences in displacement and velocity of MD were analyzed using repeated measures analysis of variance (ANOVA). The effect of episodes was not significant in the displacement levels of mothers' head MD (SS = 0.059, $F(2,46) = 1.556$, $p > .05$), but was in that of infants' (SS = 0.627, $F(2,46) = 8.672$, $p < 0.001$). A series of post-hoc pairwise t-tests were carried out as follow-up tests of the significant episode effect for infant. Consonant with the results reported in Hammal et al. (2015), significant differences in MD displacements were found between the FF and the SF, and between the SF and the RE (Table 3).

There was also a significant SFP episode effect on both the mothers' and infants' velocities in head movements (for mothers: SS = 0.048, $F(2,46) = 4.145$, $p < 0.03$; for infants: SS = 0.246, $F(2,46) = 25.434$, $p < 0.001$). Post-hoc pairwise t-tests indicated significant difference between the FF and the SF in mothers' velocities, as well as significant differences between the FF and the SF, and between the SF and the RE in infants' velocities (Table 3). These results from preliminary analyses suggest that the combined, one-dimensional MD measure of overall head movement is able to capture most of the between-episode differences previously found in the three dimensions of pitch, yaw, and roll separately.

Results from state-space modeling of TV-VAR

The bivariate VAR model previously introduced in Equations 2 were fit to the processed data. Apart from the VAR(1) model elaborated previously, we also fit a VAR model of order 2 (VAR(2)) given the previously found significant lag-2 partial correlations, which in addition to the lag-1 variables of mother and infant movements, also included the lag-2 variables in the model. We compared the fit of these two models using the Akaike Information Criterion (AIC; Akaike, 1998) and Bayesian Information Criterion (BIC; Schwarz, 1978). A lower score on either criterion suggests better model fit. VAR(2) had an AIC slightly smaller than that of VAR (1) (62812.84 vs. 62827.38) but a larger BIC (62914.08 vs 62897.47). Given AIC's tendency to prefer more complicated models, we decided to proceed with VAR(1) for model parsimony. Then, TVPs were included to capture targeted between-episode and between-dyad differences during the SFP. To help decide whether certain TVPs were supported, we began by fitting two unconditional models in which either AR or CR parameters were estimated as TVPs but not predicted using any

Table 3. Results from post-hoc pairwise t-tests.

Comparison	t-score	p-value	t-score	p-value
		Displacement		
	Infants		Mothers	
FF – SF	1.98	0.18	−3.30	0.01
FF – RE	0.39	1.00	1.33	0.59
SF – RE	−1.18	0.75	3.35	0.01
		Velocity		
	Infants		Mothers	
FF – SF	3.40	0.01	−6.51	0.00
FF – RE	1.25	0.67	−0.29	1.00
SF – RE	−1.42	0.50	5.49	0.00

covariates or assumed to take on a particular shape (Equation 3 with $f(.) = int_{ar_b}$)[4]. Figure 4 shows the estimated trajectories from the unconditional time-varying AR and CR parameters for two sample dyads. Results from fitting these unconditional models suggested that only the autoregression parameters (i.e., ar_b, and ar_m), but not the cross-regression parameters (i.e., cr_{mb} and cr_{bm}) showed evidence as TVPs, diagnosed based on the statistical significance of their corresponding process noise variances. This is also evident in plots (b1) to (b4) in Figure 4, that the estimated time-varying CR trajectories from the state-space approach were extremely flat. Therefore, CR parameters were included in the final model as time-invariant parameters.

Next, for the parameters determined to show substantial within-person, over-time variations, we added selected covariates to examine whether these over-time variations were associated with SFP episode and dyadic characteristics. In particular, we included as predictors: (1) episodic information (contrast coded into SF: SF (2) vs. others (-1) and RE: RE (1) vs. FF (-1)); (2) contextual information about the interactions: whether the mother was smiling at the moment ("$momSmile_{it}$"); and (3) and dyad-specific characteristic that may contribute to differences in dynamics: infant later attachment security ("$Richter_i$"). We also included person-specific variances of the head movement measures across all episodes (denoted as "$Var_{b,i}$" and "$Var_{m,i}$") to account for other sources of between-person variability in head movements not captured by these three sets of covariates of interest.

$$ar_{b,it} = \beta_{b0} + \beta_{b1} \times SF_{it} + \beta_{b2} \times RE_{it} + \beta_{b3}$$
$$\times momSmile_{it} + \beta_{b4} \times Richter_i + \beta_{b34} \times momSmile_{it}$$
$$\times Richter_i + \beta_{b5} \times Var_{b,i} + \zeta_{ar_b,i,t}, \text{ and} \quad (14)$$

$$ar_{m,it} = \beta_{m0} + \beta_{m1} \times SF_{it} + \beta_{m2} \times RE_{it} + \beta_{m3}$$
$$\times momSmile_{it} + \beta_{m4} \times Richter_i + \beta_{m34} \times momSmile_{it}$$
$$\times Richter_i + \beta_{m5} \times Var_{m,i} + \zeta_{ar_m,i,t}. \quad (15)$$

$$\begin{bmatrix} \zeta_{ar_b,i,t} \\ \zeta_{ar_m,i,t} \end{bmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} \psi_{ar_b,i,t} & \\ 0 & \psi_{ar_m,i,t} \end{bmatrix}\right) \quad (16)$$

Table 4 shows the estimated values for all the time-invariant parameters in Equations 2 and 14-16 along with the standard errors and the associated 95% confidence intervals of these estimates. The unit root tests were ran again with the residuals from this model, and the tests indicated the data were conditionally stationary given the model. Although the CR paramters were found to be time-invariant, their magnitudes were significantly different from zero. That is, we found evidence for mother → infant as well as infant → mother interactive contingencies during the SFP, as averaged across all three SFP episodes. That is, overall, across all dyads and all episodes, mothers' head movements at the previous time point (at a time lag of 1) were found to negatively influence children's current head movements ($cr_{mb} = -0.011$), and children's head movements at the previous time point positively influenced their mothers' ($cr_{bm} = 0.012$) at the current time point. Thus, on average across the three SFP episodes, mothers' head movement magnitudes, which we postulated to be related to their affect arousal levels, appeared to synchronize to infants' previous head movements, with higher magnitudes of infant head movements at time t-1 leading to higher magnitudes of mother head movements at time t. In contrast, the negative mother → infant interactive contingency weight suggested the opposite patterns: low magnitudes of mother head movement at time t-1 tended to elicit high magnitudes of infant head movement at time t, and conversely, high magnitudes of mother head movement at time t-1 tended to elicit low magnitudes of infant head movement at time t. These differences in interactive contingency weights may reflect mothers' intrinsic motivation to adapt to their child's head movements, and corresponding effects of the mothers on the infants either in helping to downregulate intense head movements, or in eliciting more intense movements when high magnitudes of head movement were reciprocated with low magnitudes of head movements from the others.

When we prepared data for analyses, linear trends were removed for every dyad within each episode, but all standardization was done on the non-episode-specific group level (all mothers as a group and all infants as another) instead of on the individual dyad and episode level. We made such decision for
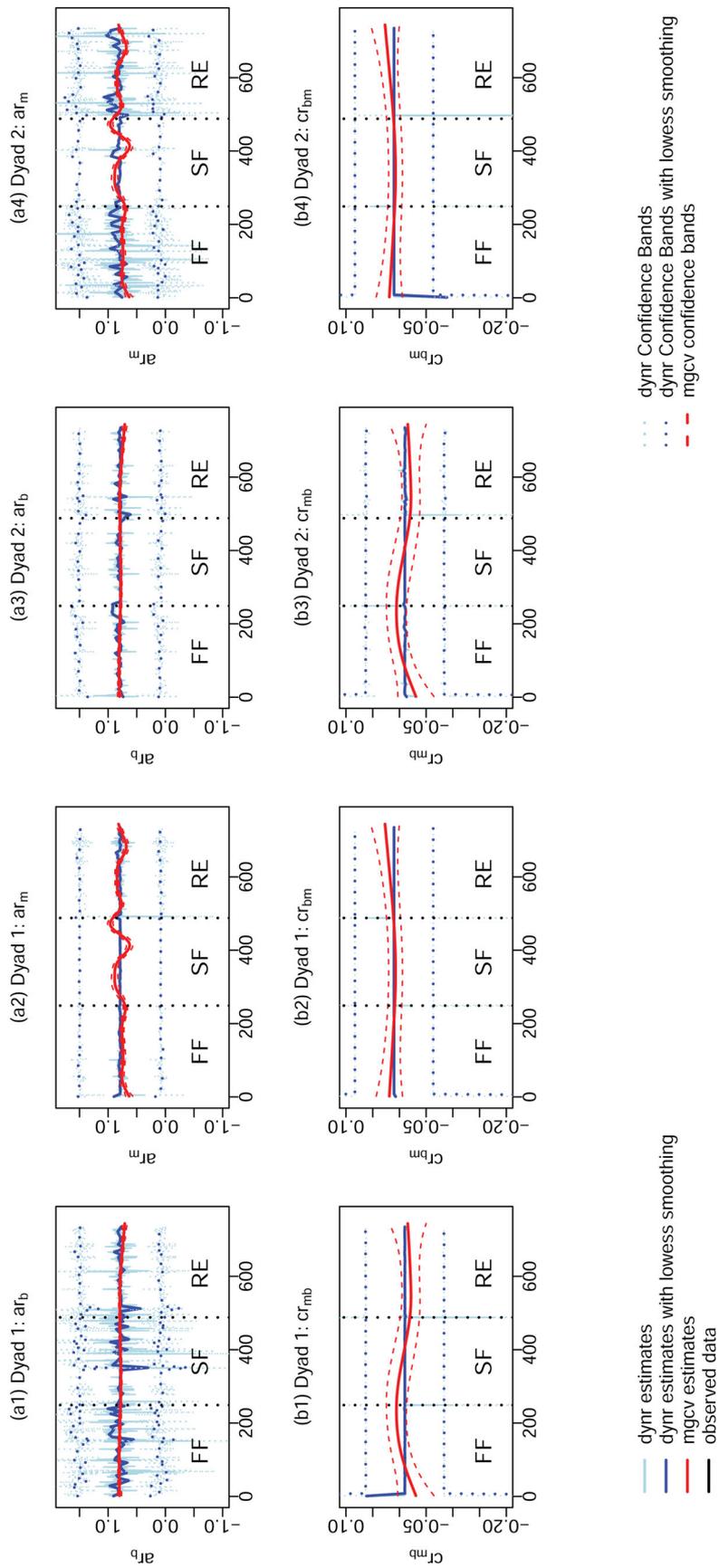
Figure 4. Comparison of AR and CR Trajectories with the Unconditional TVP Models and GAM for Two Example Dyads.

Table 4. TV-VAR model parameter estimation results using the state-space approach.

	Estimate	Standard Error	95% Confidence Interval
cr_{mb}	−0.011	0.005	(−0.020, −0.002)
cr_{bm}	0.012	0.004	(0.004, 0.020)
int_b	−0.067	0.014	(−0.095, −0.039)
int_m	0.034	0.013	(0.009, 0.059)
β_{b0}	0.730	0.013	(0.705, 0.755)
β_{b1}	−0.005	0.005	(−0.015, 0.005)
β_{b2}	−0.017	0.008	(−0.032, −0.002)
β_{b3}	−0.017	0.019	(−0.054, 0.019)
β_{b4}	0.012	0.004	(0.003, 0.020)
β_{b34}	0.008	0.006	(−0.004, 0.020)
β_{b5}	0.027	0.009	(0.009, 0.045)
β_{m0}	0.708	0.018	(0.674, 0.743)
β_{m1}	0.031	0.006	(0.020, 0.042)
β_{m2}	0.035	0.007	(0.021, 0.049)
β_{m3}	−0.002	0.016	(−0.033, 0.030)
β_{m4}	0.006	0.004	(−0.002, 0.014)
β_{m34}	−0.018	0.006	(−0.030, −0.006)
β_{m5}	0.084	0.012	(0.060, 0.107)
ψ_b	0.273	0.004	(0.265, 0.281)
ψ_m	0.219	0.003	(0.212, 0.225)
ψ_{bm}	0.006	0.002	(0.002, 0.011)
ψ_{ar_b}	0.128	0.006	(0.117, 0.139)
ψ_{ar_m}	0.144	0.006	(0.132, 0.157)
AIC: 58986.19		BIC: 59165.30	

standardization to preserve some between-dyad differences and within-dyad variability across episodes given the theoretically different natures of these three episodes of the SFP. We expected that the TVPs, namely, the time-varying AR coefficients, would be adequate in capturing most of the within-person variability across episodes. However, the intercept parameters in the TV-VAR model was estimated to be significant, albeit small in magnitude (Table 4; $int_b = -0.067, int_m = 0.034$). The significant intercepts suggested that some signs of misspecification of the VAR process noise covariance structure remained and were manifested through the intercepts.

### Within-dyad variations in AR parameters across episodes

Results showed that on average, mothers and infants manifested relatively large positive values of AR parameters, or inertia, in head movements during the SFP interactions ($\beta_{b0} = 0.730, \beta_{m0} = 0.708$). The episode-specific components in the model for time-varying AR parameters supported the experimental manipulation in SFP as different episodes resulted in different interaction dynamics indexed by head movements. As shown in Figure 5, across all dyads, infants showed lower levels of AR during the SF episode as compared to FF, though the coefficient associated with the contrast code for SF was not found significant ($\beta_{b1} = -0.005$). This might be related to the construction of the contrast code for the SF effect, in which the FF episode was grouped together with the RE episode. However, contrary to our initial expectation, there was greater similarity between infants' AR values during the RE and those from the SF, as opposed to those from the FF episode ($\beta_{b2} = -0.017$). This may reflect infants' recovery from distress developed during SF (sometimes referred to as the "carryover" effect of SF; Haley & Stansbury, 2003). A lower AR generally means that the observed process is less predictable from previous observations. Here, it reflects the infants making less consistent movements in SF and RE as compared to FF.

Mothers on average showed the highest AR values during SF ($\beta_{m1} = 0.031$). An increase in the AR parameter indicates that the movement is more predictable and consistent in time, and in the case of the SF episode, it reflects the experimental design of mothers being not responsive, thus making consistently no or minimal head movements. Mothers on average also showed higher levels of AR during the RE compared to FF ($\beta_{m2} = 0.035$), thus providing some evidence for the carryover AR effect from the SF. Still, as can be observed from Figure 5, mothers' AR was on average lower in the RE than in the SF, possibly reflecting mothers' efforts to resume her emotional connections with the infants through a variety of movements to calm and/or distract the infants. In addition, the estimated process noise variances of these two parameters, ψ_{ar_b} and ψ_{ar_m}, remained statistically significant after the inclusion of these covariates. This indicated that there was still substantial between-individual and across-time variability in these two AR parameters that were not explained by the covariates.

Between-dyad differences in inertia based on attachment

We found that the AR parameter for infants also differed depending on infants' levels of attachment security. After controlling for differences between episodes, infants who were more securely attached had higher inertia in their head movements across all episodes ($\beta_{b4} = 0.012$), which means that their head movement were more predictable compared to those less securely attached. These results are consonant with Beebe et al. (2010)'s emotional engagement results in which future securely attached infants exhibited higher levels of self-regulation and lower levels of responsivity to mother than future insecurely attached infants, as well as Jaffe et al.'s similar (2001) findings in the vocal coordination domain. Infants' secure attachment did not predict mothers' AR parameter, but it did have a significant interaction effect with mother smiling on mothers' AR parameter

AR Relation with Smile and Attachment

Episodic AR Differences
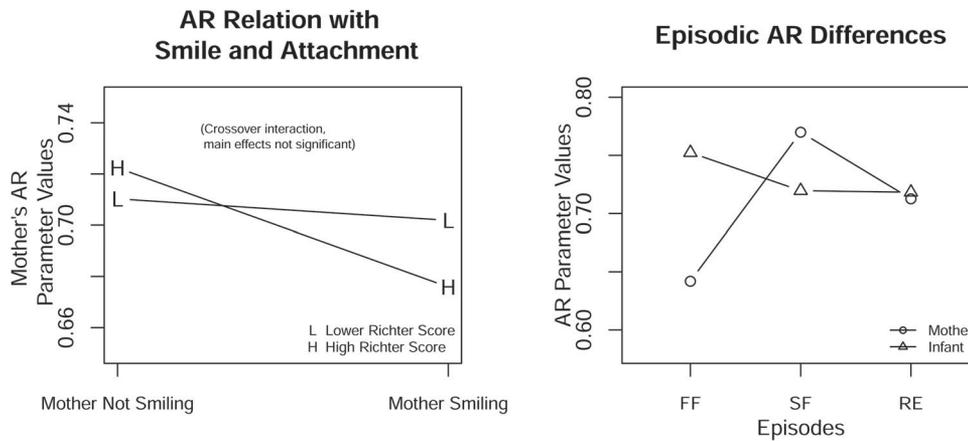
Figure 5. Differences in AR parameters dependent on whether mother was smiling, infant's attachment tendency, and episodes.

($\beta_{m34} = -0.018$), despite mother smiling itself not having any direct influence on either mothers' or infants' AR ($\beta_{b3} = -0.017$, $\beta_{m3} = -0.002$). Mothers whose infants were more securely attached had lower inertia in their head movements when they smiled compared to when they did not (see Figure 5). This suggests that they were more likely to show a greater range of head motions when they were smiling. On the other hand, mothers whose infants were less securely attached showed comparable inertia in their head movements regardless of whether they were smiling. These findings suggested that mothers of more securely attached infants exhibited greater within-person differences in their head movement when expressing positive affect–which might signal their involvement in and receptivity to the interaction –than mothers of less securely attached infants.

Results from GAM fitting of TV-VAR

We fit the model represented in Equation 12 to the same dataset with 24 dyads, over a handful of choices of number of basis functions (k) in thin plate regression splines. Figure 6 shows the estimated smooth functions for each parameter under all the k values. The choice of k made the biggest difference in mother's AR parameter (ar_m), some difference in mother's intercept parameter (int_m), and almost no difference in the other parameters. Comparing smoothed functions of difference parameters, it appeared that mother's AR parameter had the highest variability across time while the CR parameters were mostly flat (Figure 6). These observations were consistent with the conclusions from the state-space model approach. In addition, according to the plot, mother's intercept parameter may also be varying across time. The

patterns of change in mother's AR and intercept parameters also matched roughly to the change points of episodes in SFP (around time points 249 and 488).

The random sample based diagnostic in *gam.check()* did not support k being big enough for all the k values we implemented. However, the estimated trajectories suggested that wigglier functions in most of the parameters had been smoothed out with the optimized λ, and for ar_m that did exhibit a difference based on k, $k = 160$ already yielded fairly wiggly trajectories. Given this, we stopped doubling k and chose a k value based on examination of the plots. In the following comparison between the state-space model and GAM approaches, we adopt a k value of 20, which generated some wiggliness but at the same time also preserved a certain level of smoothness.

To begin with a relatively fair comparison of modeling approaches, we compared results from the two unconditional TVP models used in the state-space model approach as a model building step with results from fitting GAMs where certain parts of the model in Equation 12 were constrained to be parametric. For the unconditional model with time-varying ARs, the intercept and CR parameters (f_1, f_2, f_4 and f_5) were estimated as constants instead of functions of time, and the same goes for the intercept and AR parameters (f_1, f_2, f_3 and f_6) for the unconditional model with time-varying CRs. Estimation results for the parametric coefficients are reported in Table 5, and the estimated trajectories for TVPs are compared for two dyads in Figure 4. We would like to note that despite the TVPs not having any theoretically guided form, the models compared under the state-space approach and the GAM approach were still not entirely identical. Other than having different functional forms, the two TVPs in the state-space model approach were modeled
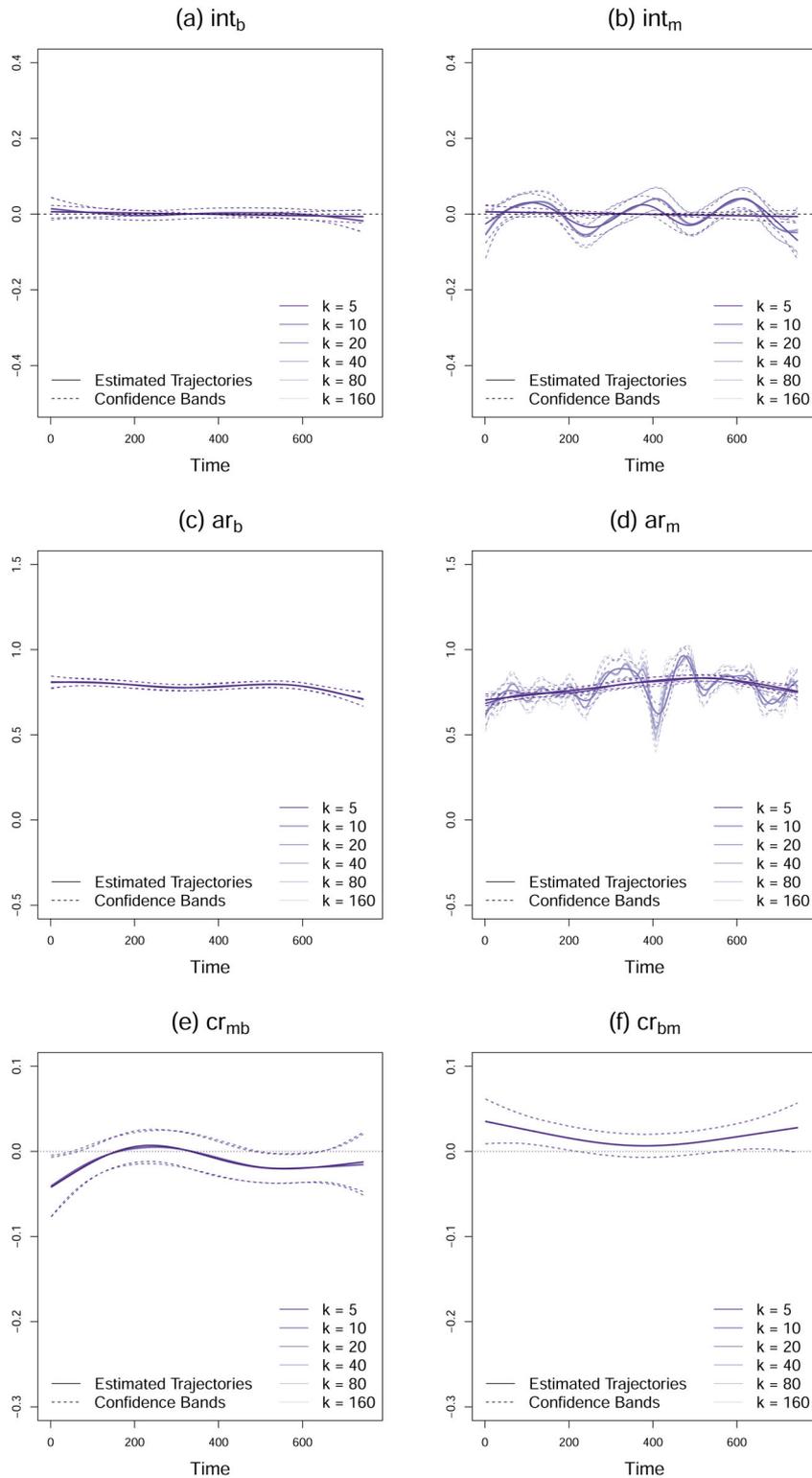
Figure 6. Time-varying Parameter Trajectories and Associated 95% Confidence Intervals Using Different Number of Basis Functions (k).

as a bivariate process, whereas in GAM the two functions representing TVPs were not associated. Despite the differences, estimates of the parametric coefficients for the ARs (in a model with time-varying CRs) and CRs

(in a model with time-varying ARs) were similar in magnitude and direction. There existed some discrepancies in the estimates of intercepts. The estimated TVP trajectories exhibited similar overall tendencies, with the

Table 5. Unconditional TV-VAR model parametric coefficients comparison between the state-space and GAM approaches.

	State-space Approach	GAM Approach
Unconditional CR Model		
cr_{mb}	−0.011*	−0.012*
cr_{bm}	0.011*	0.016*
int_b	−0.076*	0.010*
int_m	0.042*	−0.002
Unconditional AR Model		
ar_b	0.778*	0.785*
ar_m	0.796*	0.783*
int_b	0.043*	0.010*
int_m	0.040*	−0.001

*$\alpha = 0.05$.

ones from the state-space model approach being more "jiggly" due to the inclusion of random process noise (Figure 4).

The state-space model approach did not support the CR parameters as TVPs, but the time-invariance CR parameters were significantly different from zero with $cr_{mb} = -0.011$ and $cr_{bm} = 0.012$ in the final TV-VAR model (Equations 2 and 14-16). In contrast, although the smooth functions representing the time-varying CR parameters were estimated to be significant in GAM, the estimated time-varying CR trajectories were of small magnitude (similar to the results from the state-space model with unconditional time-varying CRs) and only some portions of the corresponding confidence bands for the CR trajectories did not include 0 in plots (e)-(f) in Figure 6. The mean values of the trajectories mirror these time-invariant estimates (mean($cr_{mb,t}$) = −0.011, mean($cr_{bm,t}$) = 0.015). Some discrepancies were observed in the estimated intercept parameters and the covariance structure for process noise of the head movement variables. The function representing time-varying intercept for infants was not significant in the GAM approach, and the estimated process noise variances and covariance were slightly larger compared to those in the state-space approach ($\psi_b = 0.394$, $\psi_m = 0.353$, $\psi_{bm} = 0.020$). The GAM approach also suggested the intercept parameter of mother's head movement to be time-varying, in contrast to the time-invariant intercepts we imposed in our state-space model.

To further facilitate the comparison between the state-space model with TVPs in Equations 14-16 versus the GAM model of TVPs, estimated trajectories of the shared TVP in both models, ar_b and ar_m, are plotted for two dyads in Figure 7. Comparing between dyads, the state-space approach, as discussed earlier, was able to accommodate some between-dyad differences. The smoothed ar_b trajectory from the GAM approach was roughly the mean trend in the smoothed ar_b trajectories from the state-space approach. Furthermore, the smoothed ar_m trajectories from the state-space approach show a visible SF effect in both dyads, where the mother's AR during SF (the middle chunk) was much flatter and also slightly higher in value on average than the other two episodes. This offers a demonstration of how theory-guided model can be helpful. The SF effect in mothers' AR is supported by the experiment design of SPF. In the GAM TVP trajectories, the differences between episodes were less salient but some changes in mothers' AR were still evident. On the other hand, it is worth noting that the GAM model with enforced smoothness yielded much narrower confidence bands of the estimated trajectories compared to the state-space model due to the differences between the two approaches in handling group-based models. The individual filtering and smoothing scheme adopted in the state-space model approach resulted in confidence bands reflecting the uncertainty around the estimates of the specific dyad in each plot. Meanwhile, confidence bands in the GAM approach generated through model-implied predictions reflected uncertainly at the mean sample level across all dyads.

In summary, in our particular TV-VAR model, the GAM and the state-space model approach results suggested similar overall dynamics in mother-infant interactions in terms of the estimated (mean) values of the AR/CR parameters, and both approaches uncovered some degrees of SF effect in mothers' head movements. GAM accommodated more TVPs compared to the state-space approach (6 vs. 2) but the three parameters other than AR and mother's intercept exhibited relatively flat trajectories that hovered closely around zero. In contrast, through the state-space model, we were able to link dyadic head movements to mothers' facial cue of smiling and also infants' later attachment development. These are some examples of parametric effects that are relatively straightforward to specify within the state-space framework, but are difficult to specify within the GAM framework due to software-related constraints.

Discussion

As longitudinal designs and data become more prominent in the study of human behaviors, models with TVPs have also gained considerable traction over the last decade. In this article, we applied a time series-inspired dynamical systems model with TVPs to study parent-infant co-regulation using automated measures of head movements during the SFP. We compared results from fitting variations of the TV-VAR model
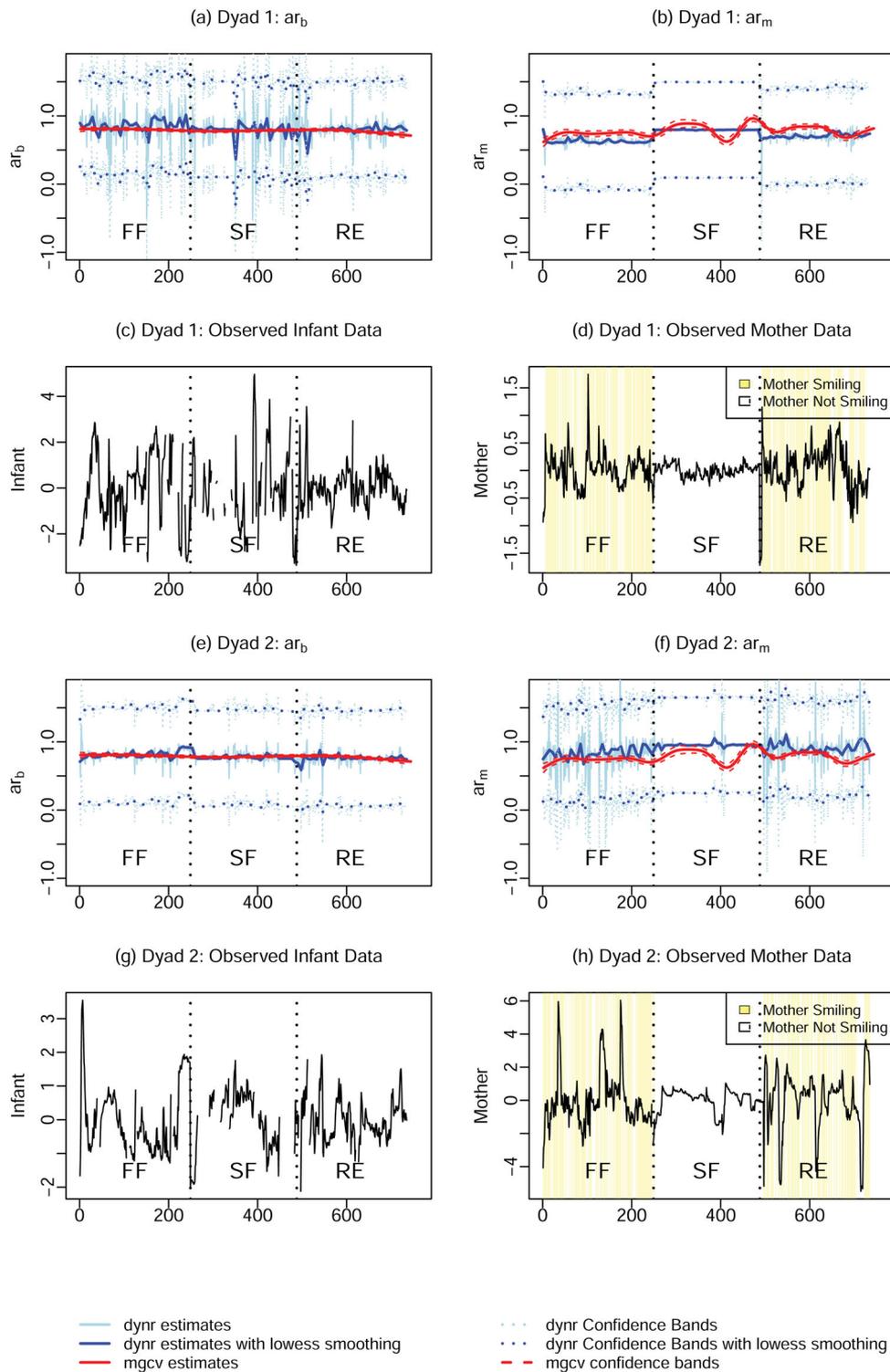
Figure 7. Comparison of AR Trajectories with the Parametric TV-VAR Model and GAM for Two Example Dyads Along with Observed Data used for Modeling.

using two approaches: a parametric state-space approach, and a semi-parametric approach utilizing GAMs. The two approaches yielded similar inferential results with regard to the mother-infant dyads'

dynamics as a group, but also some discrepancies in the findings concerning the nature of the TVPs.

Both approaches can accommodate TVPs rather flexibly, with the GAM approach being a

nonparametric approach in the TVP part, and the state-space approach accounting for additional within- and across-individual variability in the TVPs via stochastic process noises in the TVPs. Overall, the two approaches differ in both their model formulation and estimation details. Thus, the slight divergence in results did not come as a surprise, but did suggest important implications for future studies involving TVPs. To begin with, the smoothing procedure embedded in GAM with penalized regression splines allows researchers to explicitly control the smoothness/wiggliness of TVP trajectories through the smoothing parameter (λ), if desired. In the state-space modeling approach, the filtering and smoothing procedures are done individually for each unit of analysis so the smoothed TVP trajectories are also individually adapted to observed data. As such, compared to the smoothed TVP trajectories from GAM, the estimated TVP trajectories from the state-space approach tend to appear "rougher" when process noises are included to allow nuanced fluctuations in each individual's observed data be captured as process noises in the TVPs. Thus, the inclusion of process noises in the TVPs in the state-space approach, and the explicit regularization (i.e., smoothing) of the TVP trajectories through the penalty term in the GAM approach are two key features that set these approaches apart from each other.

The roughness of the TVP trajectories in the state-space approach is contributed in part by other reasons as well. Specifically, the state-space modeling approach is able to accommodate both within- and between-unit (e.g., individuals and dyads) differences in the TVPs essentially as latent variable scores, even though model fitting is performed at the group level, with time-invariant parameters that are constrained to be equal across units. In contrast, fitting a group-based model in the GAM approach is sensitive to only the universal trends and effects across all units. In principle, it is possible to adapt the TV-VAR models implemented in the two modeling approaches to obtain a model that is more comparable across frameworks. For instance, to confine the TVPs in the state-space model approach to be the same across unit (by, for example, dropping the process noise structure and adopting a functional form that is not dependent on unit-specific characteristics), or to do individual model-fitting in both approaches. Here, we can demonstrate with two dyads that some between-dyad differences exist in the AR trajectories, examining the results from both approaches with model fitting to each individual dyad (Figure 8). Alternatively, it is

also possible to specify some penalized spline functions as a state-space model (e.g., cubic splines; Chow & Zhang, 2008; Wahba, 1978). We did not adopt these procedures because we were interested in investigating specific, substantively motivated hypotheses in the state-space approach. In our view, some of these differences in results actually made our illustration more informative.

The two approaches are also characterized by distinct model identifiability constraints. In state-space models, one well-known model identifiability condition is for the system of interest to be *observable*. In other words, the system's underlying latent variable values can be uniquely determined from the observed measurements (Bar-Shalom et al., 2001). As such, the number of TVPs that can be estimated as additional latent variables in a state-space model is also limited by the number of observed endogenous (dependent) variables and latent variables that are already present in the model aside from the TVPs. Based on Gates et al. (In Press), in a model with q observed variables and p latent variables (not including TVPs), the maximum number of TVPs identifiable from the data is the minimum of p and q. Thus, in the context a VAR(1) model, for instance, the model would not be observable if more than two TVPs are present in the model. However, as distinct from the state-space approach, each smooth function in GAM that is tied to a unique predictor is identified via implicit constraints on the basis coefficients, namely, by requiring that the basis coefficients associated with any particular predictor to sum to zero over all possible values of that predictor. Because lag-1 mother and infant head movement variables are included in GAM as predictors that are distinct from the dependent variables (the lag-0 mother and infant head movement variables), we were able to allow for time-dependent smoths of the AR as well as CR parameters with the same data. However, uninformed expansion of the order of the VAR model to allow higher-lag coefficients (e.g., lag-2, lag-3 and so on) to be time-varying may yield over-fitting and is thus not recommended. In practice, we recommend that researchers first use some screening procedures, such as some of the ones we adopted in this article, to examine evidence for TVPs prior to freeing them up in completely unsupervised ways.

In summary, the two approaches have unique benefits that may make them more appealing in some cases than others. If one wants to explicitly model multivariate associations among system variables and with other TVPs, or to link the change in system
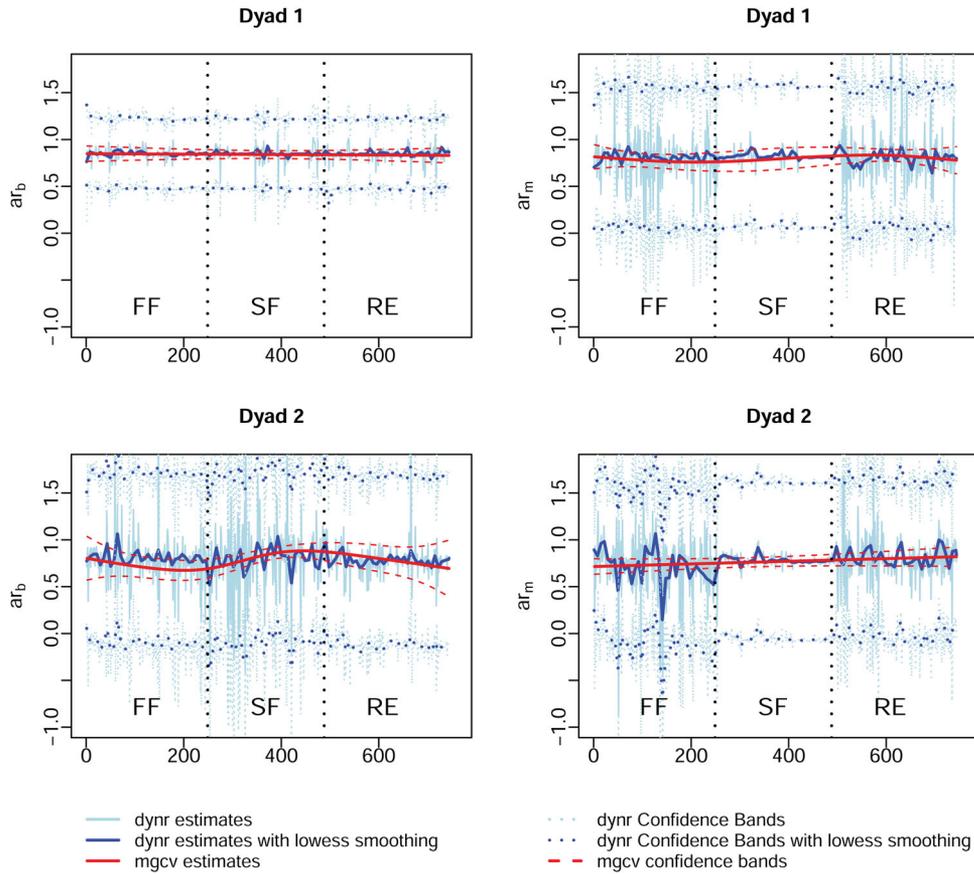
Figure 8. TVP Trajectories for Two Dyads when Fitting the State-space Model and GAM to Each Individual Dyad.

patterns (manifested in TVPs) with between-unit and within-unit characteristics to study the mechanisms of change, then the state-space model offers a more straightforward approach to implement selected parametric functions that maybe difficult or even impossible to implement and interpret in GAM. The state-space model approach can also accommodate a measurement structure, which is not possible in GAM. On the other hand, in the absence of clear parametric modeling goals, the GAM approach offers another advantage besides the flexibility associated with a model-free approach: the dependent variable of interest can conform to any distribution from the exponential family, including the Normal, Binomial, and Poisson distributions as special cases. This opens up the possibility of using binary and count data as system variables. In addition, we note that GAM through *mgcv* is only one spline-based functional regression method out of many that could be applied to estimate models with TVPs. Other functional regression software packages exist and may utilize slightly different spline formulations, penalty terms and optimization criteria (e.g. as defined within a least squares vs. likelihood framework). For example, the SAS Macro TVEM (Li et al., 2015) utilizes P-splines

and B-splines, and has extensions to accommodate data nesting, clustering and different sampling weights. Other R packages include *funreg* (Dziak et al., 2019), *npmlda* (Wu & Tian, 2018) and *refund* (Goldsmith et al., 2019). Direct comparisons of these other spline-based approaches for fitting models with TVPs are beyond of the scope of this article, but warrant further attention in a future study.

Both approaches in this study utilized a group-based model-fitting strategy, and between-unit variations were controlled by including dyad-specific characteristics into the state-space model, and allowing for stochastic system noises. A way to completely separate between-unit and within-unit variations is to fit a multilevel TV-VAR model, for which between-unit variations are modeled by random effects. *mgcv* does have the option to include random effects; it is also possible in the state-space approach by *dynr* to insert and estimate selected random effects as part of the latent variable vector concurrently with the other TVPs. However, this expands the dimension of the latent variables very quickly and brings with it other identification issues. Bayesian methods may be a more viable alternative in this case. For example, the dynamic structural equation modeling toolbox in the software

Mplus (e.g. Asparouhov et al., 2018; Hamaker et al., 2018), which is implemented through Bayesian estimation, or the R package *ctsem* (e.g. Driver & Voelkle, 2018), has the option of using *Stan* (Carpenter et al., 2017) to fit multilevel dynamic models. These extensions warrant close examination in future work.

The trajectories of the VAR processes investigated in this study (i.e., the mother and infant trajectories) were assumed to vary smoothly over time (i.e., differentiable with respect to the latent variables). Although not the focus of this article, under certain occassions the key processes of interest may display non-continuous, abrupt changes and the corresponding TVP patterns. Some approaches that would account for such kind of changes include (but are not limited to) regime-switching models (e.g. Chow et al., 2018), threshold AR models (e.g. De Haan-Rietdijk et al., 2016; Hamaker et al., 2009), and using a two-step process of change point detection followed by a parametric model with change point entered as known data.

Beyond the methodological insights discussed thus far, the article also offers an example of how dynamic models can be applied to investigate the temporal evolution of regulatory behaviors, and how TVPs can be utilized to capture within-dyad variations and between-dyad differences. To our knowledge, this article is the first effort to apply dynamic systems technique to examine co-regulation in mother-infant dyads via automated head movement measures. In confirmation with experimental manipulation and similar to the results obtained in Hammal et al. (2015), we detected some differences in mothers' head movement dynamics in the SF than in other episodes, as revealed in mothers' AR parameter. This suggests that the inertia of mothers' head movements, as revealed by their AR parameter, may serve as a proxy for understanding key changes in mother-infant intrinsic and interactive dynamics during the SFP. The decline in infants' AR parameter in the SF episode – though not statistically significant due to our a priori choice of contrast coding scheme, and the prolonged decline into the RE episode, are all in accordance with the established results in the literature on increased negative emotions during SF, the persistence of such emotions into RE (e.g. Toda & Fogel, 1993; Tronick et al., 1978). This article also linked infants' later attachment and the early interactive patterns with their mothers by showing that infants' dynamics in interaction and also their mothers' display of facial affect at 4 months are associated with 15-month attachment measure.

This article highlights the utility of using automated head movements in the study of parent-infant interactions to understand the communication and co-regulation patterns between infants and mothers. This automated measure is reasonable in cost, unaffected by the subjective biases from human raters, and allows quantification of dynamic movement patterns on a frame-by-frame basis. An ambiguity is whether infant head movement is more associated with the valence or arousal dimension of emotion. Hammal et al. (2015) found that infants' head movement was greater during tasks intended to elicit negative emotion and was strongly related to observer ratings of affect intensity. Caution needs to be exercised in interpreting our analytic results involving head movements as indices of emotion. The context within which the interactions take place may strongly influence whether head movement is more closely associated with valence or arousal. In our study, such frame-by-frame quantification also made the raw data fairly noisy. As described in Data Descriptions and Preprocessing under the Motivating Example, the raw data went through four steps of preprocessing before analyses: aggregation, variable combination, detrending within dyad and episode, and standardization with group. The decisions on whether and how to perform each step were based on the phenomenon and research question. With different research questions, the data can be preprocessed in different ways to fulfill the need. For example, in our study, aggregation was done to collapse the sampling rate within the time-scale of our phenomenon of interest – behavioral coregulation. If the researcher's interest was on a finer time-scale (e.g. simple movement tendencies for a single person), s/he can choose to not aggregate or aggregate into smaller intervals, and vice versa. With a similar dataset, a researcher can choose to detrend or not based on whether the main level change is of interest and needs to be incorporated in the modeling efforts. Lastly, whether to standardize and what level to standardize on depend on what kind of comparison the research would like to see. In our study, we standardized the data on a rather broad level given we wanted to compare both between dyads and within dyad across episodes. If the research interest is to extract common patterns of within-dyad dynamics and less about between-dyad differences, standardization can be done within each dyad instead.

One limitation of the current study is its small sample size. The analytic sample size was reduced to ensure we had enough data per dyad that spanned all three SFP episodes. Therefore, dyads in which either member had long chunks of missing data in their head movements were removed from the analytic sample. A

common reason for such missing data was that the individual moved out of range of the video camera, or that the individual's face was at a position where the software tracker could not identify the face anchors. Such missingness may potentially be classified as non-ignorable missingness (Little & Rubin, 2002), as the reason for removal was related to the variable of interest (head movements). Future studies should better account for, or directly incorporate modeling of the missing data patterns in the study, before generalizing the findings to other contexts and samples.

In conclusion, this article provides important insights on two of the most widely utilized methods for fitting dynamic models with TVPs — specifically, TV-VAR. Our empirical application further attests to the importance of considering the presence of TVPs, and the issue of self-organization, in the study of human dynamics. It also validated the feasibility of direct modeling of automated measures of head movements from a dynamic systems perspective to uncover aspects of parent-infant interaction and co-regulation. In addition, it provided further evidence for the link between interactive patterns early in life to the infant's later development of attachment.

Article information

Acknowledgments

References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Lawrence Erlbaum.

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected papers of Hirotugu Akaike* (pp. 199–213). Springer. https://doi.org/10.1007/978-1-4612-1694-0

Anderson, B. D. O., & Moore, J. B. (1979). *Optimal filtering*. Prentice Hall.

Ansley, C. F., & Kohn, R. (1985). Estimation, filtering, and smoothing in state space models with incompletely specified initial conditions. *The Annals of Statistics*, 13(4), 1286–1316. https://doi.org/10.1214/aos/1176349739

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

Bar-Shalom, Y., Li, X. R., & Kirubarajan, T. (2001). *Estimation with applications to tracking and navigation: Theory algorithms and software*. Wiley. https://doi.org/10.1002/0471221279

Beebe, B., Jaffe, J., Markese, S., Buck, K., Chen, H., Cohen, P., Bahrick, L., Andrews, H., & Feldstein, S. (2010). The origins of 12-month attachment: A microanalysis of 4-month mother–infant interaction. *Attachment & Human Development*, 12(1–2), 3–141. https://doi.org/10.1080/14616730903338985

Beebe, B., Messinger, D., Bahrick, L. E., Margolis, A., Buck, K. A., & Chen, H. (2016). A Systems view of mother-infant face-to-face communication. *Developmental Psychology*, 52(4), 556–571. https://doi.org/10.1037/a0040085

Bertenthal, B. (2007). Dynamical systems: It is about time! In S. Boker & M. Wenger (Eds.), *Data analytic techniques for dynamical systems* (pp. 1-24). Lawrence Erlbaum Associates, Inc.

Bosma, H. A., & Kunnen, E. S. (Eds.). (2011). *Identity and emotion: Development through self-organization*. Cambridge University Press. https://doi.org/10.1017/CBO9780511598425

Bringmann, L. F., Ferrer, E., Hamaker, E. L., Borsboom, D., & Tuerlinckx, F. (2018). Modeling nonstationary emotion dynamics in dyads using a time-varying vector-autoregressive model. *Multivariate Behavioral Research*, 53(3), 293–222. https://doi.org/10.1080/00273171.2018.1439722

Bringmann, L. F., Hamaker, E. L., Vigo, D. E., Aubert, A., Borsboom, D., & Tuerlinckx, F. (2017). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. *Psychological Methods*, 22(3), 409–425. http://doi.apa.org/getdoi.cfm?doi=10.1037/met0000085

Calkins, S. D. (2011). Caregiving as coregulation: Psychobiological processes and child functioning. In A. Booth, S. M. McHale, & N. S. Landale (Eds.), *Biosocial foundations of family processes* (pp. 49–59). Springer. https://doi.org/10.1007/978-1-4419-7361-0_3

Cao, J., Huang, J. Z., & Wu, H. (2012). Penalized nonlinear least squares estimation of time-varying parameters in ordinary differential equations. *Journal of Computational and Graphical Statistics*, 21(1), 42–56. http://pubs.amstat.org/doi/abs/10.1198/jcgs.2011.10021

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 1–32. https://doi.org/10.18637/jss.v076.i01

Chen, M., Chow, S.-M., & Hunter, M. D. (2018). Stochastic differential equation models with time-varying parameters. In K. van Montford, J. H. L. Oud, & M. C. Voelkle (Eds.), *Continuous-time modeling in the behavioral and related sciences* (pp. 205–238). Springer International Publishing. https://doi.org/10.1007/978-3-319-77219-6_9

Chow, S.-M. (2019). Practical tools and guidelines for exploring and fitting linear and nonlinear dynamical systems models. *Multivariate Behavioral Research*, 54(5), 690–718. (PMID: 30950646) https://doi.org/10.1080/00273171.2019.1566050

Chow, S.-M., Ferrer, E., & Nesselroade, J. R. (2007). An unscented kalman filter approach to the estimation of nonlinear dynamical systems models. *Multivariate Behavioral Research*, 42(2), 283–321. https://doi.org/10.1080/00273170701360423

Chow, S.-M., Haltigan, J. D., & Messinger, D. S. (2010). Dynamic infant-parent affect coupling during the face-to-face and still-face paradigm: Inter- and intra-dyad differences. *Emotion*, 10(1), 101–114. https://doi.org/10.1037/a0017824

Chow, S.-M., Hamaker, E. L., Fujita, F., & Boker, S. M. (2009). Representing time-varying cyclic dynamics using multiple-subject state-space models. *British Journal of Mathematical and Statistical Psychology*, 62(3), 683–716. https://doi.org/10.1037/a0017824 https://doi.org/10.1348/000711008X384080

Chow, S.-M., Ho, M. H. R., Hamaker, E. L., & Dolan, C. V. (2010). Equivalence and differences between structural equation modeling and state-space modeling techniques. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(2), 303–332. https://doi.org/10.1080/10705511003661553

Chow, S.-M., Ou, L., Ciptadi, A., Prince, E. B., You, D., Hunter, M. D., Rehg, J. M., Rozga, A., & Messinger, D. S. (2018). Representing sudden shifts in intensive dyadic interaction data using differential equation models with regime switching. *Psychometrika*, 83(2), 476–510. https://doi.org/10.1007/s11336-018-9605-1

Chow, S.-M., & Zhang, G. (2008). Continuous-time modeling of irregularly spaced panel data using a cubic spline model. *Statistica Neerlandica*, 62(1), 131–154. https://doi.org/10.1111/j.1467-9574.2007.00379.x

Chow, S.-M., Zu, J., Shifren, K., & Zhang, G. (2011). Dynamic Factor Analysis Models With Time-Varying Parameters. *Multivariate Behavioral Research*, 46(2), 303–339. http://www.tandfonline.com/doi/abs/10.1080/00273171.2011.563697

Cohn, J. F., & Tronick, E. Z. (1988). Mother-infant face-to-face interaction: Influence is bidirectional and unrelated to periodic cycles in either partner's behavior. *Developmental Psychology*, 24(3), 386–392. https://doi.org/10.1037/0012-1649.24.3.386

Cox, M., Nuevo-Chiquero, J., Saragih, J. M., & Lucey, S. (2013). *CSIRO face analysis SDK* [Paper presentation]. 10th IEEE International Conference on Automatic Face & Gesture Recognition, Shangai, China.

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

Del Negro, M., & Otrok, C. (2008). *Dynamic factor models with time-varying parameters: Measuring changes in international business cycles* (Staff Reports No. 326). Federal Reserve Bank of New York.

Driver, C. C., & Voelkle, M. C. (2018). Hierarchical bayesian continuous time dynamic modeling. *Psychological Methods*, 23(4), 774–799. https://doi.org/10.1037/met0000168

Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23(2), 283–292. https://doi.org/10.1037/h0033031

Dziak, J. J., Coffman, D. L., Reimherr, M., Petrovich, J., Li, R., Shiffman, S., & Shiyko, M. P. (2019). Scalar-on-function regression for predicting distal outcomes from intensively gathered longitudinal data: Interpretability for applied scientists. *Statistics Surveys*, 13(0), 150–180. https://doi.org/10.1214/19-SS126

Ekman, P., & Friesen, W. V. (1974). Detecting deception from the body or face. *Journal of Personality and Social Psychology*, 29(3), 288–298. https://doi.org/10.1037/h0036006

Feldman, R. (2003). Infant-mother and infant-father synchrony: The coregulation of positive arousal. *Infant Mental Health Journal*, 24(1), 1–23. https://doi.org/10.1002/imhj.10041

Gates, K., Molenaar, P. C. M., & Chow, S.-M. (In Press). *Analysis of intra-individual variation*. Taylor & Francis.

Girard, J. M., Cohn, J. F., & Torre, F. D. l. (2015). Estimating smile intensity: A better way. *Pattern Recognition Letters*, 66, 13–21. https://doi.org/10.1016/j.patrec.2014.10.004

Goldsmith, J., Scheipl, F., Huang, L., Wrobel, J., Di, C., Gellar, J. (2019). Refund: Regression with functional data [Computer software manual]. https://CRAN.R-project.org/package=refund (R package version 0.1-21)

Haley, D. W., & Stansbury, K. (2003). Infant stress and parent responsiveness: Regulation of physiology and behavior during still-face and reunion. *Child Development*, 74(5), 1534–1546. https://doi.org/10.1111/1467-8624.00621

Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models

for the affective measurements from the cogito study. *Multivariate Behavioral Research*, *53*(6), 820–841. https://doi.org/10.1080/00273171.2018.1446819

Hamaker, E. L., Zhang, Z., & Maas, H. L. J. v d. (2009). Using threshold autoregressive models to study dyadic interactions. *Psychometrika*, *74*(4), 727–745. https://doi.org/10.1007/s11336-009-9113-4

Hammal, Z., Cohn, J. F., & George, D. T. (2014). Interpersonal coordination of headmotion in distressed couples. *IEEE Transactions on Affective Computing*, *5*(2), 155–167. https://doi.org/10.1109/TAFFC.2014.2326408

Hammal, Z., Cohn, J. F., Heike, C., & Speltz, M. L. (2015). Automatic measurement of head and facial movement for analysis and detection of infants' positive and negative affect. *Frontiers in Ict*, *2*, 21. https://doi.org/10.3389/fict.2015.00021

Hammal, Z., Cohn, J. F., & Messinger, D. S. (2015). Head movement dynamics during play and perturbed mother-infant interaction. *IEEE Transactions on Affective Computing*, *6*(4), 361–121. https://doi.org/10.1007/128 https://doi.org/10.1109/TAFFC.2015.2422702

Harvey, A. C. (2001). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.

Hastie, T., & Tibshirani, R. (1993). Varying-coefficient models. *Journal of the Royal Statistical Society: Series B (Methodological))*, *55*(4), 757–796.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development*, *66*(2), 1–132. https://doi.org/10.1111/1540-5834.00136

Jeni, L. A., Cohn, J. F., & Kanade, T. (2017). Dense 3D face alignment from 2D video for real-time use. *Image and Vision Computing*, *58*, 13–24. https://doi.org/10.1016/j.imavis.2016.05.009

Johnson, S. G. (2014). *The nlopt nonlinear-optimization package*. http://ab-initio. mit. edu/nlopt.

Jokinen, K., Nishida, M., Yamamoto, S. (2010). On eye-gaze and turn-taking. In *Proceedings of the 2010 Workshop on Eye Gaze in Intelligent Human Machine Interaction* (pp. 118–123).

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, *82*(1), 35–45. https://doi.org/10.1115/1.3662552

Kelso, S. (1995). How nature handles complexity. In *Dynamic patterns: The self-organization of brain and behavior* (pp.1-28). The MIT Press.

Kleinsmith, A., & Bianchi-Berthouze, N. (2013). Affective body expression perception and recognition: A survey. *IEEE Transactions on Affective Computing*, *4*(1), 15–33. https://doi.org/10.1109/T-AFFC.2012.16

Kopp, C. B. (1982). Antecedents of self-regulation: A developmental perspective. *Developmental Psychology*, *18*(2), 199–214.

Kraft, D. (1994). Algorithm 733: Tomp–fortran modules for optimal control calculations. *ACM Transactions on Mathematical Software (Toms))*, *20*(3), 262–281. https://doi.org/10.1145/192115.192124

Kraft, D. (1988). 7). *A software package for sequential quadratic programming* (Technical Report DFVLR-FB 88-28). Institut für Dynamik der Flugsysteme.

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

Lewis, M. D., & Ferrari, M. (2001). Cognitive-emotional self-organization in personality development and personal identity. In H. A. Bosma & E. S. Kunnen (Eds.), *Identity and emotion: Development through self-organization* (p. 177–198). Cambridge University Press. https://doi.org/10.1017/CBO9780511598425

Liang, H., Miao, H., & Wu, H. (2010). Estimation of constant and time-varying dynamic parameters of hiv infection in a nonlinear differential equation model. *The Annals of Applied Statistics*, *4*(1), 460–483. https://doi.org/10.1214/09-AOAS290

Li, R., Dziak, J., Tan, X., Huang, L., Wagner, A., & Yang, J. (2015). *Tvem (time-varying effect modeling) sas macro users' guide (version 3.1.0)*. The Methodology Center, Penn State.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. Wiley.

Li, R., Tan, X., Huang, L., Wagner, A. T., Yang, J. (2014). *TVEM (time-varying effect model) SAS macro suite users' guide (version 2.1.1) [Computer software manual]*. http://methodology.psu.edu

Magnusson, D., & Cairns, R. B., (1996). Developmental science: Toward a unified framework. In R. B. Cairns, G. H. Elder Jr, & E. J. Costello, (Eds.), *Developmental science* (pp. 7–30). Cambridge University Press.

Marra, G., & Wood, S. N. (2012). Coverage properties of confidence intervals for generalized additive model components. *Scandinavian Journal of Statistics*, *39*(1), 53–74. https://doi.org/10.1111/j.1467-9469.2011.00760.x

McKeown, G. J., & Sneddon, I. (2014). Modeling continuous self-report measures of perceived emotion using generalized additive mixed models. *Psychological Methods*, *19*(1), 155–174. https://doi.org/10.1037/a0034282

Michel, G. F., Camras, L. A., & Sullivan, J. (1992). Infant interest expressions as coordinative motor structures. *Infant Behavior and Development*, *15*(3), 347–358. https://doi.org/10.1016/0163-6383(92)80004-E

Minagawa-Kawai, Y., Matsuoka, S., Dan, I., Naoi, N., Nakamura, K., & Kojima, S. (2009). Prefrontal activation associated with social attachment: Facial-emotion recognition in mothers and infants. *Cerebral Cortex*, *19*(2), 284–292. https://doi.org/10.1093/cercor/bhn081

Molenaar, P. C. M. (1987). Dynamic factor analysis in the frequency domain: Causal modeling of multivariate psychophysiological time series. *Multivariate Behavioral Research*, *22*(3), 329–353. https://doi.org/10.1207/s15327906mbr2203_5

Molenaar, P. C. M. (1994). Dynamic factor analysis of psychophysiological signals. In J. R. Jennings, P. K. Ackles, & M. G. H. Coles (Eds.), *Advances in psychophysiology: A research annual* (Vol. 5, pp. 229–302). Jessica Kingsley Publishers. https://doi.org/10.1037/003849

Molenaar, P. C. M. (2019). Granger causality testing with intensive longitudinal data. *Prevention Science*, *20*(3), 442–451. https://doi.org/10.1037/003849 https://doi.org/10.1007/s11121-018-0919-0

Molenaar, P. C. M., Gooijer, J. G. d., & Schmitz, B. (1992). Dynamic factor analysis of nonstationary multivariate

time series. *Psychometrika*, 57(3), 333–349. https://doi.org/10.1007/BF02295422

Molenaar, P. C. M., Sinclair, K. O., Rovine, M. J., Ram, N., & Corneal, S. E. (2009). Analyzing developmental processes on an individual level using nonstationary time series modeling. *Developmental Psychology*, 45(1), 260–271. https://doi.org/10.1037/a0014170

Newell, A. (1990). *Unified theories of cognition*. Harvard University Press.

Nowak, A., & Lewenstein, M. (1994). Dynamical systems: A tool for social psychology? In R. R. Vallacher & A. Nowak (Eds.), *Dynamical systems in social psychology* (pp. 17–53). Academic Press.

Ou, L., Hunter, M. D., & Chow, S.-M. (2019). What's for dynr: A package for linear and nonlinear dynamic modeling in R. *The R Journal*, 11(1), 91–111. https://doi.org/10.32614/RJ-2019-012

Ou, L., Hunter, M. D., Chow, S.-M. (2018). *dynr: Dynamic modeling in r [Computer software manual]*. https://CRAN.R-project.org/package=dynr (R package version 0.1.12-5)

Prado, R., West, M., & Krystal, A. D. (2001). Multichannel electroencephalographic analyses via dynamic regression models with time-varying lag-lead structure. *Journal of the Royal Statistical Society: Series C (Applied Statistics))*, 50(1), 95–109. http://www.blackwell-synergy.com/doi/abs/10.1111/1467-9876.00222

R Core Team. (2018). *R: A language and environment for statistical computing [Computer software manual]*. https://www.R-project.org/

Rajan, J. J., & Rayner, P. J. W. (1996). Generalized feature extraction for time-varying autoregressive models. *IEEE Transactions on Signal Processing*, 44(10), 2498–2507. https://doi.org/10.1109/78.539034

Richters, J. E., Waters, E., & Vaughn, B. E. (1988). Empirical classification of infant-mother relationships from interactive behavior and crying during reunion. *Child Development*, 59(2), 512–522. https://doi.org/10.2307/1130329

Rothbart, M. K., Ziaie, H., & O'Boyle, C. G. (1992). Self-regulation and emotion in infancy. *New Directions for Child and Adolescent Development*, 1992(55), 7–23. https://doi.org/10.1002/cd.23219925503

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464. https://doi.org/10.1214/aos/1176344136

Schweppe, F. C. (1965). Evaluation of likelihood functions for gaussian signals. *IEEE Transactions on Information Theory*, 11(1), 61–70. https://doi.org/10.1109/TIT.1965.1053737

Shiyko, M., Naab, P., Shiffman, S., & Li, R. (2014). Modeling complexity of ema data: Time-varying lagged effects of negative affect on smoking urges for subgroups of nicotine addiction. *Nicotine & Tobacco Research*, 16(Suppl 2), S144–S150. https://doi.org/10.1093/ntr/ntt109

Tarvainen, M. P., Georgiadis, S. D., Ranta–Aho, P. O., & Karjalainen, P. A. (2006). Time–varying analysis of heart rate variability signals with Kalman smoother algorithm. *Physiological Measurement*, 27(3), 225–239. https://doi.org/10.1088/0967-3334/27/3/002

Tarvainen, M. P., Hiltunen, J. K., Ranta-Aho, P. O., & Karjalainen, P. A. (2004). Estimation of nonstationary EEG with Kalman smoother approach: An application to event-related synchronization (ERS). *IEEE Transactions on Biomedical Engineering*, 51(3), 516–524. https://doi.org/10.1109/TBME.2003.821029

Toda, S., & Fogel, A. (1993). Infant response to the still-face situation at 3 and 6 months. *Developmental Psychology*, 29(3), 532–538. https://doi.org/10.1037/0012-1649.29.3.532

Tronick, E. Z., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychiatry*, 17(1), 1–13.

Turvey, M. (1990). Coordination. *American Psychologist*, 45(8), 938–953. https://doi.org/10.1037/0003-066X.45.8.938

van Geert, P. L. C. (2018). Development, complexity and dynamical systems. *International Journal of Behavioral Development Bulletin*, 1, 5–9.

Vasilenko, S. A., Piper, M. E., Lanza, S. T., Liu, X., Yang, J., & Li, R. (2014). Time-varying processes involved in smoking lapse in a randomized trial of smoking cessation therapies. *Nicotine & Tobacco Research*, 16(Suppl 2), S135–S143. doi.org/10.1093/ntr/ntt185 https://doi.org/10.1093/ntr/ntt185

Wahba, G. (1978). Improper priors, spline smoothing and the problem of guarding against model errors in regression. *Journal of the Royal Statistical Society: Series B (Methodological))*, 40(3), 364–372. https://doi.org/10.1111/j.2517-6161.1978.tb01050.x

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, 28(6), 879–896. https://doi.org/10.1002/(SICI)1099-0992(1998110)28:6<879::AID-EJSP901>3.0.CO;2-W

Wang, Q., Molenaar, P., Harsh, S., Freeman, K., Xie, J., Gold, C., Rovine, M., & Ulbrecht, J. (2014). Personalized state-space modeling of glucose dynamics for type 1 diabetes using continuously monitored glucose, insulin dose, and meal intake: An extended kalman filter approach. *Journal of Diabetes Science and Technology*, 8(2), 331–345. https://doi.org/10.1177/1932296814524080

Weiss, A. A. (1985). The stability of the AR(1) process with an AR(1) coefficient. *Journal of Time Series Analysis*, 6(3), 181–186. https://doi.org/10.1111/j.1467-9892.1985.tb00408.x

Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology))*, 65(1), 95–144. https://doi.org/10.1111/1467-9868.00374

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology))*, 73(1), 3–36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

Wood, S. N. (2013). On p-values for smooth components of an extended generalized additive model. *Biometrika*, 100(1), 221–228. https://doi.org/10.1093/biomet/ass048

Wood, S. N. (2019). *Package 'mgcv' [Computer software manual]*. https://cran.r-project.org/web/packages/mgcv/mgcv.pdf (R package version 1.8-28)

Wu, C. O., & Tian, X. (2018). *Nonparametric models for longitudinal data: With implementation in R*. Chapman and Hall/CRC.https://doi.org/10.1201/b20631

Yee, T. W. (2015). *Vector generalized linear and additive models: With an implementation in R*. Springer. https://doi.org/10.1007/978-1-4939-2818-7

Zhu, H., & Wu, H. (2007). Estimation of smooth time-varying parameters in state space models. *Journal of Computational and Graphical Statistics*, *16*(4), 813–832. http://www.jstor.org/stable/27594277 https://doi.org/10.1198/106186007X255991

Appendix

```
########### R Code Demonstration ###########

# Loading the libraries
library(dynr)# version 0.1.12-5
library(mgcv)# version 1.8-22
library(quantmod)# version 0.4-14

# The follow script needs a long-format data frame
# that should roughly looks like:
# (system variables y_baby, y_mom and other covar-
# iates omitted)

# id id.session Time.agg Time.withinSession SF RE
# 1 1.1 1    1   -1   -1
# 1 1.1 2    2   -1   -1
# .. .. ..    .. .. ..
# 1 1.2 258 1   2   0
# 1 1.2 259 2   2   0
# .. .. ..    .. .. ..
# 1 1.3 497 1  -1   1
# 1 1.3 498 2  -1   1
# .. .. .. .. ..   ..
# 2 2.1 1    1  -1  -1
# 2 2.1 2    2  -1  -1
# .. .. .. .. .. ..


#### State-space model approach: dynr —

dynrdata <- dynr.data(Data, id="id.session",
                time="Time.withinSession",
                observed=c("y_baby","y_mom"),
                covariates=c("smile","SF","RE",
                             "Richter",
                             "var_baby",
                             "var_mom"))

# Measurement model for linking latent states
# to observed variables
meas <- prep.measurement(
values.load=matrix(c(1,0,0,0,
          0,1,0,0),ncol=4,byrow=T),
params.load=matrix("fixed",ncol=4,nrow=2),
state.names=c("baby","mom","arb","arm"),
obs.names=c("y_baby","y_mom")
)

# Initial conditions for the dynamic model
initial <- prep.initial(
  values.inistate=c(0,0,.5,.5),
  params.inistate=c('fixed', 'fixed',
    'beta_b0','beta_m0'),
  values.inicov=diag(c(rep(1,2),rep(.1,2))),
  params.inicov=diag('fixed',4))
```

```
# Process noise and measurement error variances
mdcov <- prep.noise(
  values.latent=matrix(c(.5,0.1,0,0,
            0.1,.5,0,0,
            0,0,0.1,0,
            0,0,0,0.1), ncol = 4,byrow=T),
  params.latent=matrix(c('zv_mom','cov_bm',0,0,
            'cov_bm','zv_baby',0,0,
            0,0,'zv_arb',0,
            0,0,0,'zv_arm'), ncol=4,byrow=T),
  values.observed=diag(c(0,0)),
  params.observed=diag(c('fixed','fixed'),2))

# State-space model formula

# The final model in the article with only time-
# varying ARs is specified below:

formula=list(
  list(baby~intb+arb*(baby-intb)+crmb*(mom-intm),
    mom~intm+arm*(mom-intm)+crbm*(baby-intb),
    arb~beta_b0+beta_b1*SF+beta_b2*RE
+beta_b3*smile+
      beta_b4*Richter+beta_b34*Richter*smile
+beta_b5*var_baby,
    arm~beta_m0+beta_m1*SF+beta_m2*RE+beta_
      m3*smile+
      beta_m4*Richter+beta_m34*Richter*smile
+beta_m5*var_mom

))

# A dynr formula object with starting values for
# parameter optimization

dynm <- prep.formulaDynamics(formula=formula,
          startval=c(crmb=.03,crbm=.03,
          intb = 0.1,intm = 0.1,
          beta_b0=0.5,beta_b1=0.01,
          beta_b2=0.01,
          beta_b3=0.5,beta_b4=0.01,
          beta_b34=0.01,
          beta_m0=0.5,beta_m1=0.01,
          beta_m2=0.01,
          beta_m3=0.5, beta_m4=0.01,
          beta_m34=0.01,
          beta_b5=0.01,beta_m5=0.01

      ), isContinuousTime=FALSE)

# Combine all the model components specified above
# into one dynr model object

dynrmodel<-dynr.model(dynamics=dynm,measurement=meas,
          noise=mdcov, initial=initial,
          data=dynrdata, outfile="SSMTVP.c")

# Run the parameter optimization with filtering and
# smoothing for the states

modelRes <- dynr.cook(dynrmodel)
```

```
# Result summary
summary(modelRes)


#### GAM approach: mgcv —
# Created lag-1 variables
Data$bL = unlist(by(Data$y_baby,Data$id,Lag,k = 1))
Data$mL = unlist(by(Data$y_mom,Data$id,Lag,k = 1))

# Run a GAM with:
# time-varying intercepts: s(Time.agg)
# time-varying AR: s(Time.agg,by = bL) for y_baby, for
   example
# time-varying CR: s(Time.agg,by = mL) for y_baby, for
   example
```

```
# 20 basis functions: k = 20

gam_biv<-gam(list(y_baby~-1 + s(Time.agg,k = 20) +
s(Time.agg,by = bL,k = 20)+s(Time.agg,by = mL,k = 20),
       y_mom~-1 + s(Time.agg,k = 20) +
    s(Time.agg,by = mL,k = 20)+s(Time.agg,by = bL,k = 20)),
      family = mvn(d = 2),data = Data)


# Result summary
summary(gam_biv)

# Process noise variance-covariance
solve(crossprod(gam_biv$family$data$R))
```