PSYCHOMETRIKA—VOL. 75, NO. 4, 744–771
DECEMBER 2010
DOI: 10.1007/S11336-010-9176-2

USING STATE-SPACE MODEL WITH REGIME SWITCHING TO REPRESENT THE DYNAMICS OF FACIAL ELECTROMYOGRAPHY (EMG) DATA

MANSHU YANG

UNIVERSITY OF NOTRE DAME

SY-MIIN CHOW

UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

Facial electromyography (EMG) is a useful physiological measure for detecting subtle affective changes in real time. A time series of EMG data contains bursts of electrical activity that increase in magnitude when the pertinent facial muscles are activated. Whereas previous methods for detecting EMG activation are often based on deterministic or externally imposed thresholds, we used regime-switching models to probabilistically classify each individual's time series into latent "regimes" characterized by similar error variance and dynamic patterns. We also allowed the association between EMG signals and self-reported affect ratings to vary between regimes and found that the relationship between these two markers did in fact vary over time. The potential utility of using regime-switching models to detect activation patterns in EMG data and to summarize the temporal characteristics of EMG activities is discussed.

Key words: state-space model, regime-switching, time series, electromyography.

Facial electromyography (EMG) has been recognized as a powerful indicator of human emotions (Schwartz, 1975; Cacioppo & Petty, 1981; Cacioppo, Petty, Losch, & Kim, 1986; Dimberg, 1990; Dimberg, Thunberg, & Elmehed, 2000; Weyers, Hlberger, Hefele, & Pauli, 2006). When individuals are exposed to emotionally provoking stimuli, specific facial muscles are triggered (Dimberg, 1990). Previous research has shown that such changes in facial muscle activities are effectively captured by individuals' EMG recordings, even when the changes are too subtle to be detected by human raters (Schwartz, 1975; Dimberg, 1990; Cacioppo & Petty, 1981).

A time series of EMG data contains bursts of electrical activity that are typically magnified when an individual is faced with external perturbations (e.g., as part of a mood induction procedure). However, because data segments with bursts are very short and are interspersed with long periods of deactivation (i.e., segments without bursts), the distribution of EMG data deviates substantially from normality. To effectively represent the patterns of EMG data thus requires methods or models that can account for the heterogeneity in the variance of the time series, as well as the different change patterns evidenced during periods of activation and deactivation. In the present article, we propose using regime-switching state-space models (Hamilton, 1994; Kim & Nelson, 1999a) to classify such distinct periods into latent regimes. Each regime is characterized by its own predefined dynamic patterns, and we estimate the probability that an individual's manifested dynamics at the tth time point conform to a particular regime. In this way, the data are assumed to be normally distributed conditional on the regime in which an individual resides at a particular moment; heterogeneities in the dynamic patterns and variance of EMG data are also accounted for through the incorporation of these latent regimes.

Preparation of this article was supported in part by the National Science Foundation grant BCS-0826844 awarded to Sy-Miin Chow.

Requests for reprints should be sent to Manshu Yang, Department of Psychology, University of Notre Dame, Notre Dame, IN 46556, USA. E-mail: myang@nd.edu

The Dynamics of EMG Data

Despite the increased prevalence of EMG data in psychology, most studies involving facial EMG were restricted to using Analysis of Variance (ANOVA) or correlation analyses to identify "static" between-subject differences (Cacioppo et al., 1986; Davis et al., 1995; Robinson, Cinciripini, Carter, Lam, & Wetter, 2007; Ravaja, Saari, Puttonen, & Keltikangas Jarvinen, 2008). In those studies, real-time EMG responses were typically aggregated within a trial or a condition, and mean amplitudes over time were used for analytic purposes. As a result, the temporal information embedded in EMG data was lost.

There has been a long-standing interest in classifying EMG data into regimes or portions with similar dynamics even though the lack of suitable methodological tools has often deterred contemporary affect researchers from pursuing ideas along this line. Cacioppo, Martzke, Petty, & Tassinary (1988), for instance, categorized EMG responses over the Corrugator Supercilii (CS) region—a small pyramidal muscle region that pulls the eyebrows into a frown, the magnitude of which serves as a measure of the intensity of negative affect—into four specific forms. They include EMG cluster, EMG mound, EMG spike and baseline EMG (see Figures 1(a)–(d), respectively). Cacioppo et al. (1988) found that, in general, participants felt more negative when EMG activity was elevated rather than staying around baseline level; and EMG clusters were associated with higher ratings of negative affect than did EMG mounds and spikes. Such distinctions between different EMG activation patterns cannot be extracted unequivocally from aggregate data.

Importantly, analyzing the dynamics of EMG data in real time provides a direct glimpse into theoretically interesting aspects of individual differences in emotion regulation. Davidson (1998), for instance, argued that individual differences in emotion regulation may stem from differences in the *threshold*, *peak* or *amplitude* associated with a certain emotion. In addition, *the rise time to peak* and *the recovery time* in an emotion process may differ across individuals. This concurs with the results reported, for instance, by Gilboa and Revelle (1994). For example, people tend to focus on duration when making judgement about sadness, whereas they are more inclined to focus on magnitude, or so called peak intensity, when rating anger (Gilboa & Revelle, 1994). Information concerning individual differences in affective dynamics can thus provide key insights into the development of psychopathology (Larsen, 2000) and help make the measurement of affective styles more tractable (Davidson, 1998). These time-specific attributes can only be extracted if the temporal dynamics of EMG data are preserved. In the present article, we show how information from regime-switching models can be used to derive quantifiable, individual-specific estimates of these time-specific features.

Objective of the Present Article

Within the field of psychology, researchers have rarely evaluated the dynamics of facial EMG over time, not to mention constructing a meaningful statistical model to describe the associated changes as dynamic processes. Our key objective in the present article is to illustrate how regime-switching state-space models can be used to capture the dynamic characteristics of EMG data. In particular, we seek to represent changes in activities in individuals' Corrugator Supercilii region and their correspondence with individuals' self-reported affect intensity. Model development is guided by the need to address two practical issues pertaining to the analysis of EMG data.

The first issue we seek to address is whether and how researchers can accurately identify periods with heightened EMG activation and nonactivation. Previous methods have focused on using deterministic, or arbitrarily imposed thresholds to identify the point(s) at which an individual switches to an activation phase. For instance, one common criterion is to define an activation phase as the points at which an individual's absolute EMG level exceeds a certain number of

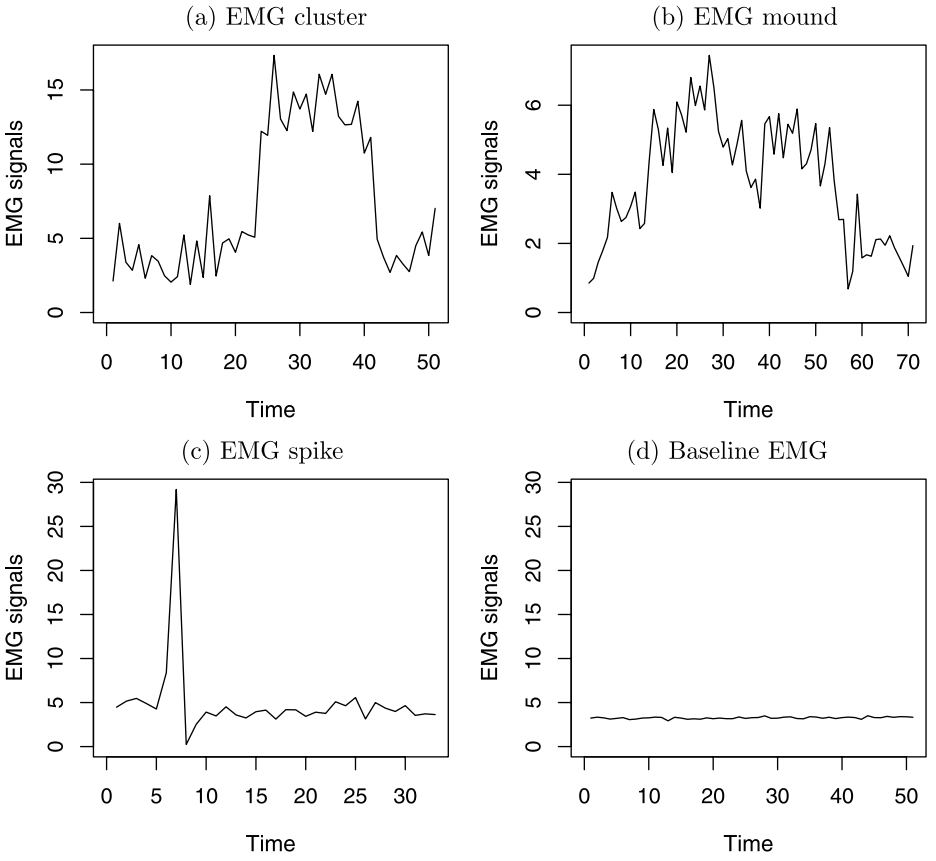

FIGURE 1.
Four Forms of EMG Dynamics: (**a**) EMG cluster, a multimodal response less than 5 seconds in duration and marked by abrupt onset and offset; (**b**) EMG mound, a relatively smooth response less than 5 seconds in duration and with gradual onset and offset; (**c**) EMG spike, a unimodal response less than 2 seconds in duration and with sharp onset and offset and (**d**) baseline EMG, a stable EMG activity lasting for at least 10 seconds in duration. Note that one unit of Time on the horizontal axes represents 0.2 second.

standard deviations (e.g., 3 SDs) relative to the individual's baseline EMG activity (for a review, see Hodges & Bui, 1996; Staude & Wolf, 1999). More often, the subjective assessments of expert coders are still used as the gold standard against which such judgments are made. More recently, Bayesian change point models have been proposed as an alternative to traditional deterministic methods to provide a more rigorous and automated way of identifying such transition or change points (Johnson, Elashoff, & Harkema, 2003). Our approach has some parallels to the Bayesian change point approach, except that we used regime-switching models to probabilistically classify each individual's time series into latent "regimes" characterized by similar error variance and dynamic patterns. In other words, we believe that it is plausible to divide a time series of EMG data into a few discrete stages, or so-called "regimes," on which subsequent model construction is based. Thus, instead of allowing for differences in dynamics after each of the change points and allowing for infinitely many change points as in Johnson et al. (2003), we only specify a few, theoretically driven regimes, defined by known characteristics of EMG data. Using the Kim filter and maximum likelihood procedures for parameter estimation (Kim & Nelson, 1999a), model fitting is characterized by substantially reduced computational costs compared to standard Markov chain Monte Carlo techniques used to estimate fully Bayesian models such as

that proposed by Johnson et al. (2003). For illustration purposes, we also compute individual-based temporal characteristics such as *rise time* and *longest duration of burst* (Davidson, 1998) using information from the regime-switching models.

The second issue of interest in the present context concerns the possible divergence and convergence between EMG and self-report data as emotion regulation unfolds. EMG responses and self-report affect ratings have been thought to reflect unique aspects of individuals' emotion regulatory process (Gilbert, 2007). For instance, compared with depressed patients, schizophrenic inpatients with flat affect were found to show reduction in emotional display (e.g., showing reduced Zygomatic or cheek EMG activity and longer pauses during dyadic interactions), but no difference in self-reported emotion intensity (Sison, Alpert, Fudge, & Stern, 1996). Stated differently, from an *inter*-individual standpoint, the relationship between psychophysiological and self-report markers of emotions is at best meager in previous studies involving data that are aggregated over time. In contrast, when EMG activities and self-reported affect ratings are both recorded in real time, the nature of their relationship and any within-person changes therein can be evaluated more systematically. For instance, at heightened arousal level, the physiological and perceptual aspects of emotions may switch from a decoupled phase to manifesting synchronous changes.

In summary, the key objective of the present article is to present a novel application of regime-switching state-space models to detect activation patterns in individuals' facial EMG data during emotion regulation processes and their correspondence with self-report data. In the next few sections, we first review our motivating example, followed by an overview of the general modeling framework within which our proposed regime-switching models are structured. Next, specifics of the models considered in our empirical application and the associated estimation procedures are outlined. Finally, we summarize the results from empirical model fitting and a simulation study. Potential promises and utility brought by this modeling approach are elaborated within the context of our modeling results.

1. Motivating Empirical Example

Four participants' data were selected from a larger emotion study and used for model-fitting purposes in this article. In the experiment, participants were subjected to an emotion induction session during which a series of 23 pictures selected from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2005) was displayed for 6 seconds each to evoke negative affect. Concurrently, miniature Ag/AgCl electrodes were attached to the participants' faces to record their facial EMG activities over the Corrugator Supercilii (CS) region. In addition, the participants were told to rate their ongoing affect intensity on a scale of 1 to 7 using a slider throughout the entire emotion induction session. The facial EMG signals and continuous self-reports were acquired using the MP150 system (Biopac Systems, Inc., 2005). The sampling rate across all channels was fixed at 1000 Hz to accommodate the fast-varying nature of our key dependent variable of interest (i.e., EMG activity). A low-pass filter was used with a cut-off frequency of 500 Hz and a high-pass filter was preset to 10 Hz.

It is worth noting that raw EMG data essentially reflect differences in electrical potential measured between recording electrodes. Thus, negative and positive values of EMG activity are of no qualitative difference but rather, indicate flow in different directions (Fridlund & Cacioppo, 1986). Thus, the raw EMG data were first rectified by taking absolute values of the signals. Subsequently, each EMG time series was integrated over every 100 samples by taking a running summation of the area under the signal curve (Biopac Systems, Inc., 2005). The resultant integrated EMG data provide a more appropriate basis for evaluating mean changes and will be used in the present study to portray changes in affect-related trends.

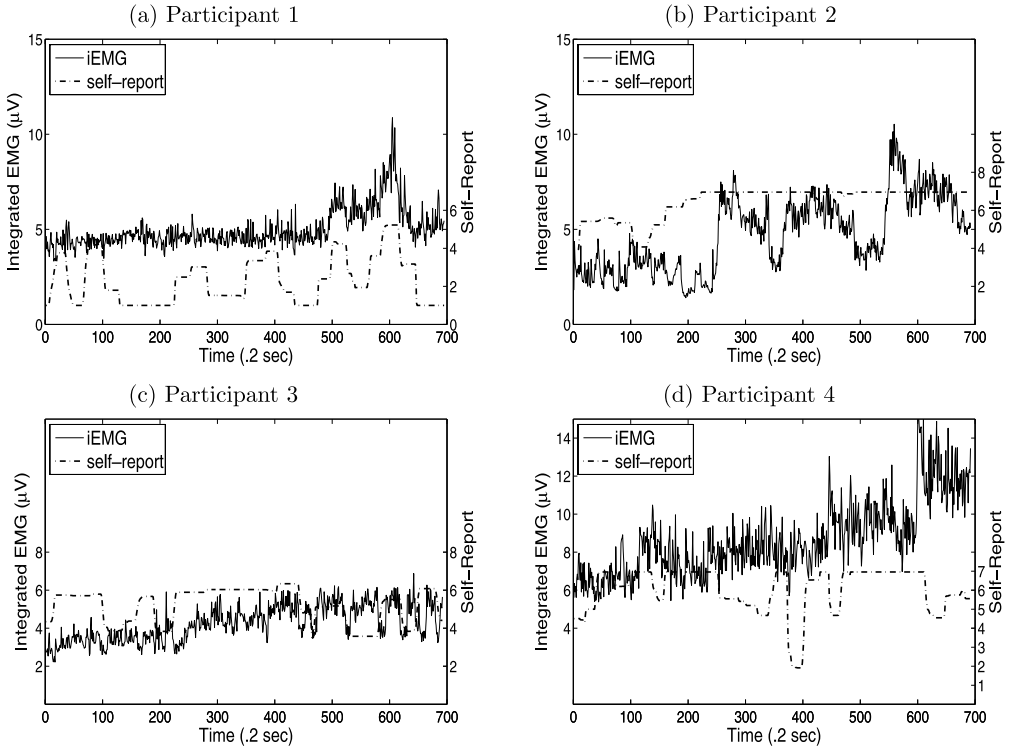

FIGURE 2.
Plots of integrated EMG and self-report affective ratings for participants 1 to 4 in panels (**a**)–(**d**), respectively. The total length of time series for each participant ranges from 692 observations to 695 observations, with a time interval of 0.2 second between two adjacent observations. Self-report = self-report affect ratings; iEMG = integrated EMG signals.

For data reduction purposes, each time series of integrated EMG data was aggregated over every 200 time points. Subsequently, a spline smoothing and interpolation procedure (Akima, 1970b, 1970a) was applied to each EMG time series to get rid of artifact spikes in the EMG data caused by body movements or electronic interferences that were irrelevant to the emotion processes under model. As noted by Fridlund and Cacioppo (1986), a rule of thumb for the choice of sampling rate is to sample 4–8 times faster than the highest frequency of interest to allow reconstruction of the original waveform with minimal smoothing. In this article, the original sampling rate was 1,000 samples/sec, so after aggregation there would still be 5 observations within each second. Since EMG responses evoked by emotional changes usually last for more than 1 second (Cacioppo et al., 1988), the aggregation process still preserved the affect-related EMG responses of interest in the present study. Data aggregation and smoothing procedure were conducted using the software of R (R Development Core Team, 2009).

The aggregated, integrated EMG signals and self-reports of Participants 1 to 4 are plotted in Figures 2(a)–(d), respectively. Visual inspection of the data reveals that the emotion induction procedure did in fact lead to heightened EMG activities in several portions of the data. For Participant 1, for instance, the deactivation phase occurred approximately during the first 500 observations, followed by a transition to an activation phase between the 500th and the 695th observation, and a return to a deactivation phase around the 695th observation. Note that an activation phase is characterized by elevated EMG signals with relatively large variance, whereas a deactivation phase is characterized by "flat," baseline fluctuations in small amplitude. In addition, within each participant, the association between EMG and self-reports was observed to

vary over time. For instance, during the first 500 observations, EMG signals were found to show rapid, more-or-less random baseline fluctuations in spite of the structured changes in his/her self-reports. Starting approximately from the 500th observation, the ebbs and flows in the participant's integrated EMG began to mirror the change patterns of his/her self-reports. Thus, preliminary graphical explorations confirmed our conjecture that a time-varying relationship may in fact exist between EMG and self-report data.

In sum, the distinct dynamics of EMG activities during the activation and deactivation phases suggest that a two-regime model provides a reasonable starting point for grouping EMG data into meaningful theoretical phases. The emergence of a synchronous relationship between self-reports and EMG signals during heightened EMG activation also points to the tendency for the association between self-reports and EMG signals to strengthen during the activation phase. Due to potential timing differences in the transition between regimes and possible individual differences in the onset and offset of EMG activity, we chose to perform model fitting at the individual level. This kind of idiographic-oriented approach is frequently adopted in studies involving the analysis of real-time psychophysiological data due to the availability of intensive repeated measurement data from each individual. Deriving individual-based estimates also enables more accurate prediction and classification at the individual level.

2. Regime-Switching State-Space Models

Regime-switching state-space models are a useful modeling tool in situations where a system evolves through time in a discontinuous fashion, or when certain parameters or characteristics (i.e., mean and variance) of the system switch between several discrete "stages," or namely, regimes. The state-space modeling framework is composed of two equations: a *measurement equation* and a *dynamic equation* (Harvey, 1989; Durbin & Koopman, 2001). The *measurement equation* describes the relationship between a set of observed variables and a set of latent variables over time. The *dynamic equation* describes the dynamics of the latent variables, namely, how latent variables change over time. A state-space model with regime-switching parameters can be expressed as

$$\mathbf{y}_t = \mathbf{d}_{S_t} + \mathbf{H}_{S_t}\boldsymbol{\alpha}_t + \mathbf{A}_{S_t}\mathbf{x}_t + \boldsymbol{\epsilon}_t, \tag{1}$$

$$\boldsymbol{\alpha}_t = \mathbf{c}_{S_t} + \mathbf{F}_{S_t}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t, \tag{2}$$

$$\begin{bmatrix} \boldsymbol{\epsilon}_t \\ \boldsymbol{\eta}_t \end{bmatrix} \sim N\left(0, \begin{bmatrix} \mathbf{R}_{S_t} & 0 \\ 0 & \mathbf{Q}_{S_t} \end{bmatrix}\right), \tag{3}$$

where Equations (1) and (2) are the *measurement equation* and *dynamic equation*, respectively. To ease presentation, subject index is omitted from the modeling equations because it is assumed that all the modeling components, including parameters and latent variables, are allowed to be different across individuals. In Equations (1) and (2), $\mathbf{y}_t$ is an $n \times 1$ vector of variables observed at time t; $\boldsymbol{\alpha}_t$ is a $k \times 1$ vector of unobserved latent variables at time t; $\mathbf{H}_{S_t}$ is an $n \times k$ factor loading matrix that links the observed variables $\mathbf{y}_t$ to the vector of latent variables $\boldsymbol{\alpha}_t$; $\mathbf{d}_{S_t}$ is an $n \times 1$ vector of intercepts; $\mathbf{x}_t$ is an $r \times 1$ vector of exogenous variables observed at time t; $\mathbf{A}_{S_t}$ is an $n \times r$ matrix which links the exogenous predictors to the observed data; $\boldsymbol{\epsilon}_t$ is an $n \times 1$ vector of measurement errors assumed to be serially uncorrelated over time and normally distributed with a mean vector of zeros and covariance matrix $\mathbf{R}_{S_t}$ at each time point t; $\mathbf{c}_{S_t}$ is a $k \times 1$ vector of intercepts and $\mathbf{F}_{S_t}$ is a $k \times k$ transition matrix that links the latent variables vector $\boldsymbol{\alpha}$ at time t to the latent variables at the previous time point; $\boldsymbol{\eta}_t$ is a $k \times 1$ vector of dynamic errors at time t, which are serially uncorrelated and normally distributed with zero means and covariance

matrix $\mathbf{Q}_{S_t}$. Note that the subscript S_t in the vectors/matrices $\mathbf{d}_{S_t}$, $\mathbf{H}_{S_t}$, $\mathbf{A}_{S_t}$, $\mathbf{c}_{S_t}$, $\mathbf{F}_{S_t}$, $\mathbf{R}_{S_t}$, and $\mathbf{Q}_{S_t}$ indicates that some parameters in them may be regime-switching and their values thus depend on the latent, discrete-valued regime indicator, S_t.

State-space models with regime-switching parameters provide a suitable tool for representing EMG dynamics because the activation and deactivation phases embedded in EMG signals can be readily regarded as two discrete "regimes" in a regime-switching model. In addition, the time-varying association between EMG and self-reports can be mathematically modeled by treating the regression slope of EMG signals on self-reports as a regime-dependent parameter in a state-space model. To make inferences on S_t, it is essential to specify a *transition probability* matrix, which indicates the probability that an individual is in a certain regime conditional on the previous regime. In matrix form, these transition probabilities are written as

$$\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MM} \end{bmatrix}, \tag{4}$$

where $\Pr[S_t = j | S_{t-1} = i] = p_{ij}$, $i, j = 1, 2, \ldots, M$ and $\sum_{j=1}^{M} p_{ij} = 1$ (Kim & Nelson, 1999a). The parameter p_{ij} represents the probability that an individual is in the jth regime at time t given that the individual is in the ith regime at time $t-1$. These transition probabilities are regarded as model parameters that are to be estimated with other time-invariant parameters in $\mathbf{d}_{S_t}$, $\mathbf{H}_{S_t}$, $\mathbf{A}_{S_t}$, $\mathbf{c}_{S_t}$, $\mathbf{F}_{S_t}$, $\mathbf{R}_{S_t}$, and $\mathbf{Q}_{S_t}$ in Equations (1) to (3).

3. Models for Empirical Application

Many dynamic models can be written in a state-space form to represent a variety of different dynamics. In the present study, three different models with increasing complexity were considered for our empirical application. Let $\mathbf{Y}_t = [y_1, y_2, \ldots, y_t]$ denote an individual's time series of EMG activities recorded at time 1 to t, and let $\mathbf{Z}_t = [z_1, z_2, \ldots, z_t]$ denote the individual's self-reported affect intensity at time 1 to t.

3.1. Model 0: Linear Regression Model with ARMA Errors (BASE Model)

Our baseline model was a linear regression model with no regime-switching, expressed as

$$\begin{aligned} y_t &= \mu_y + \beta z_t + e_t, \\ e_t &= \phi e_{t-1} + \zeta_t + \theta \zeta_{t-1}. \end{aligned} \tag{5}$$

In this model, μ_y is the intercept, β is a time-invariant regression slope and e_t is the error term at time t. Unlike the traditional linear regression model with independent errors, the residual term, e_t, is assumed to follow a first-order autoregressive moving-average (ARMA) process to account for the autocorrelation patterns observed in the EMG series.[1] The variables ζ_t and ζ_{t-1} in Equation (5) represent "shocks" at time t and $t-1$, respectively, that define the moving

[1] We expected the autocorrelations to reflect within-person changes in affect that were of substantive interest. In some cases, standard data pre-processing procedures (e.g., integration, smoothing and averaging) such as the ones adopted in the present study may also create "arbitrary" autocorrelation patterns in the data. Whatever the reasons might be, the inclusion of an ARMA(1,1) component was needed from a practical standpoint in the present context and the lag order was determined by inspecting the autocorrelation and partial autocorrelation plots of the EMG data.

average part of the ARMA(1,1) process. They are assumed to follow a normal distribution with a mean of 0 and a variance of σ_ζ^2 for $t = 1, \ldots, T$. The model in Equation (5) can be expressed in state-space form by specifying the vectors and matrices in Equations (1) and (2) as

$$\boldsymbol{\alpha}_t = \begin{bmatrix} e_t & \theta\zeta_t \end{bmatrix}', \quad \mathbf{x}_t = z_t, \quad \boldsymbol{\eta}_t = \begin{bmatrix} \zeta_t & \theta\zeta_t \end{bmatrix}',$$
$$\boldsymbol{\epsilon}_t = 0, \quad \mathbf{d} = \mu_y, \quad \mathbf{H} = \begin{bmatrix} 1 & 0 \end{bmatrix},$$
$$\mathbf{A} = \beta, \quad \mathbf{c} = \begin{bmatrix} 0 & 0 \end{bmatrix}', \quad \text{and} \quad \mathbf{F} = \begin{bmatrix} \phi & 1 \\ 0 & 0 \end{bmatrix}.$$

This model was selected as a baseline model because it provided a reasonable starting point for capturing the overall autocorrelation patterns of the participants and it served to indicate the general, time-invariant relationship between EMG and self-report data when the heterogeneity in error variance was not accounted for.

3.2. Model 1: Model with Regime-Dependent Intercept and Regression Slope (RS Model)

As shown in the motivating example section, a shift in the association between EMG and self-report responses from a decoupled to a synchronous phase is observed in individuals' EMG signals during heightened emotion activation. By allowing the regression intercept and slope, namely, μ_y and β, to be regime-dependent, or in other words, to switch between the two regimes/phases, the relationship between EMG and self-reports can be allowed to change in a discrete manner between the activation phase and the deactivation phase. The resultant model is expressed as

$$\mathbf{y}_t = \mu_{yS_t} + \beta_{S_t} z_t + e_t, \tag{6}$$
$$e_t = \phi e_{t-1} + \zeta_t + \theta\zeta_{t-1}, \tag{7}$$

where the intercept, μ_{yS_t}, and the regression slope, β_{S_t}, are allowed to be regime-dependent as

$$\mu_{yS_t} = \begin{cases} \mu_{y0} & \text{if } S_t = 0 \text{ (deactivation phase)}, \\ \mu_{y1} & \text{if } S_t = 1 \text{ (activation phase)}, \end{cases} \tag{8}$$

and

$$\beta_{S_t} = \begin{cases} \beta_0 & \text{if } S_t = 0 \text{ (deactivation phase)}, \\ \beta_1 & \text{if } S_t = 1 \text{ (activation phase)}. \end{cases} \tag{9}$$

In this model, S_t is a latent dummy variable indicating which regime is in effect, and its value at time t needs to be estimated using the Kim filter algorithm (to be described in the next section). This model was established based on preliminary examination of the correlations between EMG and self-reports throughout the trial.

3.3. Model 2: Model with Regime-Dependent AR and MA Coefficients (RS-ARMA Model)

Model 2 was constructed to extend Model 1 by allowing the autoregressive coefficient, ϕ, and the moving average coefficient, θ, to be regime-dependent. It is expressed as

$$\mathbf{y}_t = \mu_{yS_t} + \beta_{S_t} z_t + e_t, \tag{10}$$
$$e_t = \phi_{S_t} e_{t-1} + \zeta_t + \theta_{S_t}\zeta_{t-1}. \tag{11}$$

By allowing ϕ and θ to be regime-specific, we indirectly allowed the variance of the residual component, $E(e_t^2)$, to be heterogeneous across the deactivation and activation stages. Whereas

TABLE 1.
Properties of the three models considered in the empirical application.

Model	Acronym	Name	Regression slope β	AR coefficient ϕ	MA coefficient θ
0	BASE	Linear Regression Model with ARMA Errors	Time-invariant	Time-invariant	Time-invariant
1	RS	Model with Regime-Dependent Intercept and Regression Slope	Regime-specific	Time-invariant	Time-invariant
2	RS-ARMA	Model with Regime-Dependent AR and MA Coefficients	Regime-specific	Regime-specific	Regime-specific

we expected the regime-dependent intercept in Model 1 to be able to capture part of the elevations in EMG signals during the activation phase, we used the regime-dependent AR and MA parameters to better capture the more subtle differences in fluctuation patterns around this elevated intercept. For instance, the AR(1) coefficient might be close to 1.0 in portions of the data that showed trends or fluctuations in big amplitudes.

The regression intercept, μ_{yS_t}, the regression slope, β_{S_t}, the AR coefficient, ϕ_{S_t}, and the MA coefficient, θ_{S_t}, are dependent on the same dummy variable, S_t. That is, their changes are hypothesized to be dictated by the same regime indicator, S_t. Thus, the regime-dependent parameter vector, $\boldsymbol{\psi}_{S_t} = [\mu_{ySt}, \beta_{S_t}, \phi_{S_t}, \theta_{S_t}]$, is equal to $[\mu_{y0}, \beta_0, \phi_0, \theta_0]$ when $S_t = 0$ (at deactivation phase) and $[\mu_{y1}, \beta_1, \phi_1, \theta_1]$ when $S_t = 1$ (at activation phase).

The key features of the three models are summarized in Table 1. In the most complex model (Model 2), the parameter vector $\boldsymbol{\psi} = [p_{00}, p_{01}, \mu_{y0}, \mu_{y1}, \beta_0, \beta_1, \sigma_\zeta^2, \phi_0, \phi_1, \theta_0, \theta_1]$ consists of all the time-invariant parameters that were to be estimated by maximum likelihood procedures. The first two parameters are the probability of staying in the deactivation phase from time $t-1$ to time t, and the probability of switching from a deactivation phase at time $t-1$ to an activation phase at time t, respectively. Other elements of the transition probability matrix, namely, p_{10} and p_{11}, are known once these parameters have been estimated.

4. Estimation Procedures: The Kim Filter and Maximum Likelihood Procedures

The unknown elements in a state-space model can be classified into two categories: the time-varying latent variables, $\boldsymbol{\alpha}_t$ and S_t, and the time-invariant parameter vector, $\boldsymbol{\psi}$. Once a state-space regime-switching model has been structured, the Kim filter and the related Kim smoother can be used to estimate the values of the latent variables in $\boldsymbol{\alpha}_t$,[2] as well as the probability of being in a certain regime at each time point. We outline the Kim filter algorithm and refer interested readers to Kim and Nelson (1999a) for further details. The notations for all the variables in the Kim filter are listed in Table 2. It is worth noting that all the expectation and covariance values in Table 2 are inherently conditional on the parameter vector $\boldsymbol{\psi}$, although it is omitted for clarity (e.g., $E[\boldsymbol{\alpha}_t|S_t = j, \mathbf{Y}_t] \triangleq E[\boldsymbol{\alpha}_t|S_t = j, \mathbf{Y}_t, \boldsymbol{\psi}]$).

[2]These latent variable estimates are usually the true scores or factor scores associated with an underlying process of interest. In our modeling example, the only latent variable that was to be estimated was the ARMA component, e_t.

TABLE 2.
A summary of the notations used in the Kim filter.

Notation	Meaning
$\mathbf{Y}_t$	$\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_t\}$
$\boldsymbol{\alpha}_{t\|t-1}^{i,j}$	$E[\boldsymbol{\alpha}_t \| S_{t-1} = i, S_t = j, \mathbf{Y}_{t-1}]$
$\boldsymbol{\alpha}_{t\|t}^{i,j}$	$E[\boldsymbol{\alpha}_t \| S_{t-1} = i, S_t = j, \mathbf{Y}_t]$
$\mathbf{P}_{t\|t-1}^{i,j}$	$\text{Cov}[\boldsymbol{\alpha}_t \| S_{t-1} = i, S_t = j, \mathbf{Y}_{t-1}]$
$\mathbf{P}_{t\|t}^{i,j}$	$\text{Cov}[\boldsymbol{\alpha}_t \| S_{t-1} = i, S_t = j, \mathbf{Y}_t]$
$\boldsymbol{\alpha}_{t\|t}^{j}$	$E[\boldsymbol{\alpha}_t \| S_t = j, \mathbf{Y}_t]$
$\mathbf{P}_{t\|t}^{j}$	$\text{Cov}[\boldsymbol{\alpha}_t \| S_t = j, \mathbf{Y}_t]$
$\mathbf{v}_t^{i,j}$	$E[\mathbf{y}_t - \mathbf{H}_{S_t}\boldsymbol{\alpha}_t - \mathbf{A}_{S_t}\mathbf{x}_t \| S_{t-1} = i, S_t = j, \mathbf{Y}_{t-1}]$
$\mathbf{f}_t^{i,j}$	$\text{Cov}[\mathbf{y}_t - \mathbf{H}_{S_t}\boldsymbol{\alpha}_t - \mathbf{A}_{S_t}\mathbf{x}_t \| S_{t-1} = i, S_t = j, \mathbf{Y}_{t-1}]$

The Kim filter is essentially an estimation procedure that combines the traditional Kalman filter and the Hamilton filter (Kim & Nelson, 1999a). The Kalman filter provides a way to derive estimates of the latent variables in $\boldsymbol{\alpha}_t$ based on both current and previous regimes as well as the manifest observations up to time t (i.e., $E[\boldsymbol{\alpha}_t | S_{t-1} = i, S_t = j, \mathbf{Y}_t]$, denoted herein as $\boldsymbol{\alpha}_{t|t}^{i,j}$). In contrast, the Hamilton filter offers a way to update the probability of being in the jth regime at time t conditional on manifest observations up to time t (i.e., $\Pr[S_t = j | \mathbf{Y}_t]$).

The Kim filter can be implemented in three sequential steps. First, the Kalman filter is executed to obtain latent variable estimates, $\boldsymbol{\alpha}_{t|t}^{i,j}$, and their covariance matrix, $\mathbf{P}_{t|t}^{i,j}$, given both current and previous regimes. Next, the Hamilton filter is implemented to get the joint conditional regime probability of being in the jth regime at time t and the ith regime at time $t-1$, namely, $\Pr[S_{t-1} = i, S_t = j | \mathbf{Y}_t]$, as well as the regime probability of being in the jth regime at time t, namely, $\Pr[S_t = j | \mathbf{Y}_t]$. Third, a "collapsing process" (p. 10, Kim & Nelson, 1999a) is carried out using estimates from the first two steps. Given M possible regimes, the "collapsing process" allows a researcher to combine (or "collapse") the $M \times M$ sets of latent variable estimates, $\boldsymbol{\alpha}_{t|t}^{i,j}$, and their covariance matrix, $\mathbf{P}_{t|t}^{i,j}$ (with $i = 1, \ldots, M$, $j = 1, \ldots, M$), into M sets of filtered estimates, $\boldsymbol{\alpha}_{t|t}^{j}$ (i.e., $E[\boldsymbol{\alpha}_t | S_t = j, \mathbf{Y}_t]$), and the associated covariance matrix, $\mathbf{P}_{t|t}^{j}$, which are all that is needed to start the next iteration in the first step again. The whole filtering process is carried out recursively (i.e., sequentially for each time point) until the latent variable estimates, $\boldsymbol{\alpha}_{t|t}^{j}$, and regime probabilities, $\Pr[S_t = j | \mathbf{Y}_t]$, have been calculated for time $t = 1$ to T. The latter, in particular, is a key component in our subsequent derivation of individual difference measures, such as rise time.

Maximum likelihood procedures can be used to derive optimal estimates of all the time-invariant parameters in a regime-switching model. Under normality assumptions of the measurement and dynamic errors, the prediction errors, which capture the discrepancies between the manifest observations and the predictions implied by the model (see definition for the prediction errors, $\mathbf{v}_t^{i,j}$, and their covariance matrix, $\mathbf{f}_t^{i,j}$, in Table 2), are also normally distributed. They are obtained as by-products from the Kim filter. As a result, a likelihood function can then be written using these by-products as

$$f\left(\mathbf{y}_t | S_t = j,\ S_{t-1} = i, \mathbf{Y}_{t-1}\right) = (2\pi)^{-p/2}\left|\mathbf{f}_t^{i,j}\right|^{-1/2} \exp\left\{-\frac{1}{2}\left(\mathbf{v}_t^{i,j}\right)'\left(\mathbf{f}_t^{i,j}\right)^{-1}\mathbf{v}_t^{i,j}\right\}. \quad (12)$$

Equation (12) is also known as the *prediction error decomposition* function (Harvey, 1989). More formally, maximum likelihood estimates of the time-invariant parameters in $\boldsymbol{\psi}$ can be obtained by alternating between the Kim filter to yield elements for the likelihood function in Equation (12) and an optimization step (e.g., Newton–Raphson method) to maximize the loglikelihood function until some convergence criteria have been met.

If the entire time series of observations is available for estimation purposes—as is the case in most studies in psychology, one can refine the estimates of the latent variables in $\boldsymbol{\alpha}_t$ and the probability of the unobserved regime indicator, S_t, based on all the observed information in the sample, yielding the smoothed latent variable estimates, $\boldsymbol{\alpha}_{t|T} = E(\boldsymbol{\alpha}_t|\mathbf{Y}_T)$, and the smoothed regime probabilities, $\Pr[S_t = j|\mathbf{Y}_T]$. These elements can be estimated by means of the Kim smoother. Estimates from the Kim smoother, $\boldsymbol{\alpha}_{t|T}$ and $\Pr(S_t = j|\mathbf{Y}_T)$, are more accurate than those from the Kim filter, since the former is based on information from the entire sample rather than from previous information up to the current observation, as in the Kim filter. More detailed descriptions of the Kim filter, the Kim smoother and other calculation steps are included in Appendix A.

5. Empirical Results

5.1. Model Selection and Validation

The three proposed models (Models 0, 1, and 2) were fitted to the data from the four participants using MATLAB[3] (The MathWorks, Inc., 1991). Because not all of the models are nested within the most complex model, the likelihood ratio test is not appropriate here for model selection purposes. To aid model selection, the Akaike Information Criterion (AIC; Akaike, 1973) and the Bayes Information Criterion (BIC; Schwarz, 1978) were used. Given the very large number of observations in this study (e.g., approximately 700 observations for each individual time series) and the tendency for the AIC and BIC to favor more complex models in such cases, we supplemented results based on the AIC and BIC with information from the cumulative sum (CUMSUM) test (Harvey, 1981). The CUMSUM test is an exploratory method for detecting model misfit and structural changes. The null hypothesis in a CUMSUM test is that the cumulative sum of the standardized prediction errors, $\mathbf{v}_t$ (see Table 2), in a time series follows a Brownian process. By constructing significance lines based on the null hypothesis, researchers can detect substantial deviations in the cumulative prediction errors from these predefined boundaries and in doing so, detect increase in model misspecification and possible structural changes in the system (Harvey, 1989). In addition, autocorrelation and partial autocorrelation plots of the prediction errors obtained from model fitting were used to determine if there were significant autocorrelation patterns that were unaccounted for in the prediction errors. Q–Q plots of the prediction errors were also examined to ensure that the conditional normality assumption of the prediction errors was satisfied.

5.2. Computing Rise Time and Longest Burst Duration

One of the modeling goals we sought to accomplish in the context of the proposed regime-switching models was to derive actual, mathematical estimates of quantities such as *rise time* and *longest burst duration* in accordance with the theoretical definitions suggested by Davidson (1998). One way to operationalize these theoretical concepts within the context of the proposed regime-switching models is as follows. First, every time point in each individual's time series

[3] All MATLAB codes used in this study (including the Kim filter and maximum likelihood procedures) were written by the authors of this article. The codes are available upon request from the first author.

TABLE 3.
Goodness of Fit Statistics for All Participants across All Models.

	Fit indices	Model 0 (BASE)	Model 1 (RS)	Model 2 (RS-ARMA)	Final selected model
Participant 1	Deviance(-2LL)	1059	1040	1022	1028
	AIC	1069	1058	1044	1044
	BIC	1092	1099	1094	1080
Participant 2	Deviance(-2LL)	1218	1193	1172	1174
	AIC	1228	1211	1194	1192
	BIC	1251	1252	1244	1233
Participant 3	Deviance(-2LL)	979	909	876	877
	AIC	989	927	898	895
	BIC	1012	968	948	936
Participant 4	Deviance(-2LL)	1968	1930	1923	1925
	AIC	1978	1948	1945	1943
	BIC	2001	1989	1995	1984

is categorized into either an activation phase or a deactivation phase based on the individual's smoothed regime probabilities, $\Pr(S_t = j|\mathbf{Y}_T)$, obtained from the Kim smoother. That is, the individual is classified to be in an activated phase if $\Pr(S_t = 1|\mathbf{Y}_T) \geq 0.5$. In contrast, the individual is classified to be in a deactivated phase if $\Pr(S_t = 1|\mathbf{Y}_T) < 0.5$. Since the participants always began in a deactivated phase at the start of the experimental session, the rise time for an individual can be defined to be the time length between the first observation and the earliest time point at which the individual first switches to an activation phase. If a particular facial EMG region is activated for less than 2 seconds, the EMG activities within this short period are usually regarded as an EMG spike caused by artifacts, as opposed to reflecting actual affective changes (Cacioppo et al., 1988). Therefore, only activated periods that persisted for more than 2 seconds were considered as activation phases. By the same token, we defined the longest duration of burst to be the longest, continuous time length during which an individual remained in an activation phase.

5.3. Model Fitting Results

The AIC values (see Table 3) indicated that the RS-ARMA model (Model 2) was preferred over the first two models. The BIC, however, did not favor the most complex model for all participants (e.g., Participant 1; see Table 3). We proceeded to constraining parameters that were not significantly different from zero in Model 2 (determined on an individual-by-individual basis) to be zero and fitted Model 2 again to each participant's data. The resultant models, termed the "Final Selected Model" in Tables 3 and 4, yielded the lowest BIC values among all the models considered. The final selected model, namely, the simplified RS-ARMA model with individualized constraints, was thus chosen as the best-fitting model. We focus herein on elaborating the results from this model. For comparison purposes, the parameter estimates obtained from all models are summarized in Table 4.

The regression slope associated with self-reports in the "deactivation" regime (i.e., β_0) was found to be non-significantly different from zero for all participants in Model 2 and was therefore fixed at zero for all participants. For Participants 1 and 2, the lack of association between EMG activities and self-reports during the deactivation phase changed into a statistically significant positive association during the activation phase. Thus, constraining the association between facial EMG and self-reports to be constant over time as in traditional regression models (i.e., the BASE

TABLE 4.
Parameter estimates and standard errors for all four participants.

	Parameters	Estimates (Standard errors)			
		Model 0 (BASE)	Model 1 (RS)	Model 2 (RS-ARMA)	Final selected model
Participant 1	p_{00}	–	0.996 (0.003)	0.992 (0.004)	0.994 (0.004)
	p_{10}	–	0.009 (0.007)	0.018 (0.010)	0.012 (0.009)
	μ_{y0}	–	4.596 (0.122)	4.569 (0.064)	4.561 (0.028)
	μ_y/μ_{y1}	4.778 (0.362)	4.626 (0.207)	4.720 (0.168)	4.670 (0.169)
	β_0	–	−0.002 (0.050)	−0.003 (0.027)	=0
	$\beta(\beta_1)$	0.133 (0.062)	0.525 (0.064)	0.484 (0.054)	0.516 (0.051)
	ϕ_0	–	–	0.332 (0.167)	0.245 (0.052)
	$\phi(\phi_1)$	0.981 (0.008)	0.833 (0.075)	0.679 (0.113)	0.538 (0.046)
	θ_0	–	–	−0.084 (0.160)	=0
	$\theta(\theta_1)$	−0.686 (0.034)	−0.567 (0.109)	−0.302 (0.197)	=0
	σ_ζ^2	0.269 (0.014)	0.251 (0.014)	0.223 (0.012)	0.230 (0.013)
Participant 2	p_{00}	–	0.980 (0.012)	0.975 (0.010)	0.977 (0.010)
	p_{10}	–	0.154 (0.095)	0.046 (0.021)	0.052 (0.026)
	μ_{y0}	–	4.971 (2.128)	4.208 (2.256)	4.782 (0.659)
	μ_y/μ_{y1}	5.186 (2.267)	−1.502 (2.563)	−3.532 (2.630)	−2.934 (1.844)
	β_0	–	−0.027 (0.284)	0.075 (0.310)	=0
	$\beta(\beta_1)$	−0.034 (0.310)	1.098 (0.353)	1.382 (0.367)	1.301 (0.247)
	ϕ_0	–	–	0.984 (0.014)	0.985 (0.006)[a]
	$\phi(\phi_1)$	0.984 (0.007)	0.987 (0.007)	0.985 (0.009)	0.985 (0.006)[a]
	θ_0	–	–	−0.332 (0.057)	−0.338 (0.056)
	$\theta(\theta_1)$	−0.424 (0.036)	−0.446 (0.042)	−0.784 (0.073)	−0.793 (0.089)
	σ_ζ^2	0.341 (0.018)	0.285 (0.019)	0.265 (0.016)	0.267 (0.017)

model) would risk masking the more subtle shifts in relationship between the two. This was not the case for Participants 3 and 4, whose regression slopes remained non-significantly different from zero during both the activation and deactivation phases.

Compared with Model 1 (RS; with regime-invariant AR and MA parameters), the final selected model with regime-dependent AR and MA parameters yielded a better fit, thus providing evidence that there were differences in the AR coefficient, ϕ, and the MA coefficient, θ, between the activation phase and the deactivation phase. Figure 3 shows the smoothed regime probabilities of being in the activation phase $\Pr(S_t = 1|\mathbf{Y}_T)$ over time for Participants 1 to 4 based on the final selected model (see Figures 3(a)–(d)). For Participant 1, his/her AR parameter was closer to 1.0 during the activation phase, and this coincided well with the increased fluctuations in the participant's integrated EMG during the activation phase (after $t = 500$). For Participants 3 and 4, in contrast, the AR parameter was closer to 1.0 during the deactivation phase due to the existence of an upward trend in these participants' data before transitioning formally into the "activation" phase.[4] For Participant 2, no clear difference was observed in the AR parameter in the two regimes and this parameter was therefore constrained to be equal across regimes.

Clear differences existed in the intercepts of all the participants' EMG data in the two regimes, with the activation phase being associated with a higher intercept than the deactiva-

[4]For Participant 4, a slight upward trend was observed throughout the entire time series. To account for this positive trend, we also fitted an alternative model in which time was incorporated as an exogenous predictor to Participant 4's data. The results obtained were largely similar to those summarized in Table 4 and were thus omitted due to space constraints. In the context of Models 0–2, the trend in Participant 4's EMG data was captured in part by the participant's estimated AR weights, which were close to 1.0 in all models.

TABLE 4.
(*Continued*)

	Parameters	Estimates (Standard errors)			
		Model 0 (BASE)	Model 1 (RS)	Model 2 (RS-ARMA)	Final selected model
Participant 3	p_{00}	–	0.977 (0.008)	0.973 (0.008)	0.974 (0.009)
	p_{10}	–	0.047 (0.016)	0.054 (0.017)	0.052 (0.017)
	μ_{y0}	–	3.518 (0.443)	3.552 (0.413)	3.941 (0.110)
	μ_y/μ_{y1}	4.112 (0.605)	5.362 (0.489)	5.441 (0.203)	5.297 (0.043)
	β_0	–	0.060 (0.083)	0.076 (0.077)	=0
	$\beta(\beta_1)$	0.039 (0.110)	−0.029 (0.096)	−0.029 (0.039)	=0
	ϕ_0	–	–	0.940 (0.024)	0.945 (0.022)
	$\phi(\phi_1)$	0.940 (0.015)	0.935 (0.020)	0.728 (0.103)	0.707 (0.103)
	θ_0	–	–	−0.444 (0.067)	−0.460 (0.067)
	$\theta(\theta_1)$	−0.400 (0.043)	−0.592 (0.057)	−0.596 (0.129)	−0.572 (0.127)
	σ^2_ζ	0.242 (0.013)	0.171 (0.010)	0.159 (0.010)	0.161 (0.010)
Participant 4	p_{00}	–	0.999 (0.001)	0.998 (0.002)	0.998 (0.002)
	p_{10}	–	0.002 (0.003)	0.001 (0.001)	0.001 (0.001)
	μ_{y0}	–	8.668 (0.721)	8.346 (0.759)	7.891 (0.405)
	$\mu_y\mu_{y1}$	9.330 (1.063)	10.865 (1.870)	10.331 (1.205)	12.094 (0.301)
	β_0	–	−0.076 (0.097)	−0.057 (0.107)	=0
	$\beta(\beta_1)$	−0.026 (0.117)	0.287 (0.288)	0.306 (0.213)	=0
	ϕ_0	–	–	0.965 (0.013)	0.964 (0.015)
	$\phi(\phi_1)$	0.982 (0.009)	0.972 (0.011)	0.930 (0.055)	0.933 (0.050)
	θ_0	–	–	−0.691 (0.056)	−0.677 (0.057)
	$\theta(\theta_1)$	−0.658 (0.040)	−0.743 (0.041)	−0.822 (0.092)	−0.808 (0.077)
	σ^2_ζ	1.010 (0.054)	0.939 (0.050)	0.928 (0.050)	0.932 (0.050)

[a]The AR(1) weight for Participant 2 was constrained to be equal across regimes based on results from Model 2.

Note: The subscripts "0" and "1" for μ_y, β, ϕ and θ are used to identify the regime-specific parameters during the deactivation phase and the activation phase, respectively. The variances of the ARMA residual process, $E(e_t^2)$, were 0.245, 4.021, 0.515, 2.018 for Participants 1 to 4, respectively, during the deactivation phase, and 0.324, 0.598, 0.167, 1.044 for Participants 1 to 4, during the activation phase.

tion phase (except for Participant 2). All the participants were also observed to have negative MA weights. That is, when the participants were subjected to "shocks" that led to a heightened EMG level at the current time point, such shocks tended to lower the EMG level at the next time point. This gave rise to fast fluctuations in EMG level (e.g., in the form of spikes) that diminished completely after just one lag. The activation phase was characterized by a higher negative MA weight, indicating that such rapid fluctuations were more likely to occur during the activation phase.

It is of interest to understand the collective role of the AR and MA weights when considered as an integrated system. The model-implied variance of the ARMA residual process, $E(e_t^2)$, can be computed as $\frac{(\hat{\theta}_0^2+2\hat{\phi}_0\hat{\theta}_0+1)\hat{\sigma}_\zeta^2}{1-\hat{\phi}_0^2}$ and $\frac{(\hat{\theta}_1^2+2\hat{\phi}_1\hat{\theta}_1+1)\hat{\sigma}_\zeta^2}{1-\hat{\phi}_1^2}$ for the deactivation and activation phases, respectively.[5] These variance estimates are summarized as a footnote in Table 4. For Participant 1, the MA weight was not significantly different from zero in both regimes. Thus, the higher AR(1)

[5]These formulas are appropriate only if $|\phi| < 1$, which is in fact the case for the parameter estimates in the present study.

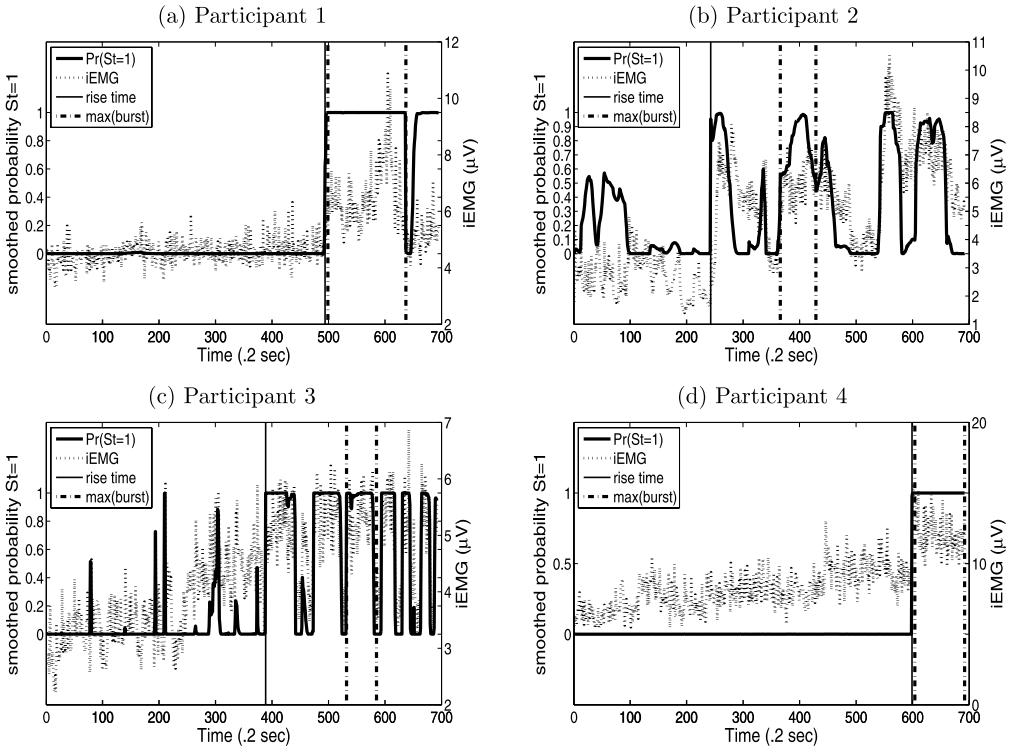

FIGURE 3.
The smoothed probabilities of being in the activation phase $\Pr(S_t = 1|Y)$ for the four participants based on the final selected model are shown in panels (**a**)–(**d**), respectively. According to the final selected model, the *longest durations of burst* were estimated to be 28.6, 12.6, 10.6 and 18.6 seconds for Participants 1 to 4, respectively; They were marked as the period between the two dashed vertical lines in panels (**a**)–(**d**). The *rise times* were estimated to be 98.8, 77.8, 42.0 and 119.8 seconds for Participants 1 to 4, respectively. They were marked as the period between the first time point and the solid vertical line in panels (**a**)–(**d**).

weight in the activated regime served to capture the increased amplitude of the participant's EMG data during the activation phase. The AR(1) also gave rise to higher model-implied residual variance in the activation than in the deactivation phase. For Participants 3 and 4, the deactivation phase was characterized by more pronounced upward trends, which were manifested as AR(1) weights that were closer to unity during the deactivation phase. Their higher (i.e., less negative) MA(1) weights during the deactivation phase were less "effective" at reversing the surges in integrated EMG levels compared with the more negative MA(1) weights observed during the activation phase. This led to higher model-implied variance estimates for the residual processes of the two participants during the deactivation than the activation phase. Participant 2 manifested largely similar dynamics, except that the participant's AR(1) weight was found to be equal across regimes and the discrepancy between the participants' MA weights in the two regimes (and consequently, the residual variance estimates) was also greater compared with other participants. In sum, by allowing for regime-specific intercept, regression slope, as well as AR and MA parameters, the final selected model was able to capture the heterogeneity in the mean and variance structures of the participants' EMG data during the activation versus the deactivation phase.

The smoothed probabilities from the final selected model were examined and subsequently used to compute the longest duration of burst and the rise time for each participant (Figure 3). Substantial individual differences were found in terms of the timing and frequency of switching between the two regimes. Within each individual, the smoothed probabilities reflected the

activation status of the EMG data well. In other words, periods with high probabilities of being in the activation phase were observed to coincide with periods with heightened EMG signals, thus providing support for the utility of the proposed regime-switching models as a classification tool. Results indicated that the longest burst duration for Participant 3 was shorter than those associated with the other three participants, which implied that Participant 3 was more likely to alternate between the activation and deactivation phases. In contrast, it took longer for Participants 1 and 4 to first rise to an activated phase compared to Participants 2 and 3.

Inspection of the smoothed probability plots described above confirmed that the final selected model provided a clear and automated way to classify the participants' data into substantively interpretable activation and deactivation phases. For the purpose of a posteriori model validation, several additional model fit indices and diagnostic methods were used to examine the performance of the final selected model, including the CUMSUM test, Q–Q plot, autocorrelation plot and partial autocorrelation plot of the prediction errors. Results from the CUMSUM test indicated that the standardized prediction errors from the final selected model lay between the 95% significance lines for all participants, indicating no substantial deviations in fit throughout the span of the individual data. Q–Q plots of the prediction errors confirmed that the normality assumption was met in the best-fitting model. In contrast, an evaluation of the Q–Q plots of the prediction errors from Model 0, for instance, revealed a mixture of normal distributions, further validating our choice of using regime-switching state-space models to represent such deviations from normality. The corresponding autocorrelation and partial autocorrelation plots indicated no statistically significant autocorrelation trends in the prediction errors of the final selected model. Higher order autoregressive or moving average components were, therefore, not needed.

6. Monte Carlo Simulation Study

6.1. Simulation Designs

To evaluate the performance of the Kim filter and the associated parameter estimation procedure, a simulation study was conducted. Simulated data were generated using the best-fitting model, namely the RS-ARMA model, and the same model was fitted to the resultant data. Different conditions were constructed with different combinations of population parameter values. The true values of parameters were selected to match the range of parameter estimates from the earlier empirical results.

Three modeling aspects were manipulated in this study. A total of 8 conditions was evaluated with

(a) the transition probabilities chosen from two sets of values:

$$\begin{bmatrix} p_{00} = 0.99 & p_{01} = 0.01 \\ p_{10} = 0.01 & p_{11} = 0.99 \end{bmatrix} \tag{13}$$

or

$$\begin{bmatrix} p_{00} = 0.90 & p_{01} = 0.10 \\ p_{10} = 0.10 & p_{11} = 0.90 \end{bmatrix}, \tag{14}$$

(b) the AR coefficient in the deactivation regime, ϕ_0, set to either $\phi_0 = 0.3$ or $\phi_0 = 0.6$; and
(c) the MA coefficient in the deactivation regime, θ_0, specified to be either $\theta_0 = -0.3$ or $\theta_0 = -0.5$.

For each condition, a total number of 100 time series, each consisting of 500 time points, were generated. For each simulated time series, the true model was fitted to the data and the associated time-invariant parameters were estimated by means of the ML method. Summary statistics of the parameter estimates across the 100 Monte Carlo runs were then computed.

TABLE 5.
Parameter estimates when $\theta_0 = -0.3$.

p_{00}	ϕ_0	Parameter	$\boldsymbol{\psi}$	$\bar{\hat{\boldsymbol{\psi}}}$	Bias	Emp. SE	Ave. SE	rBias	RMSE
$p_{00} = 0.99$	$\phi_0 = 0.3$	p_{00}	0.99	0.99	−0.003	0.01	0.01	−0.21	0.01
		p_{10}	0.01	0.01	0.001	0.01	0.01	0.12	0.01
		μ_{y0}	3.00	3.01	0.012	0.11	0.08	0.11	0.11
		μ_{y1}	4.00	3.99	−0.008	0.33	0.17	−0.02	0.33
		β_0	0.00	−0.01	−0.005	0.05	0.03	−0.10	0.05
		β_1	0.50	0.50	−0.002	0.11	0.07	−0.02	0.11
		σ_ζ^2	0.25	0.25	−0.004	0.02	0.02	−0.25	0.02
		ϕ_0	0.30	0.09	−0.209	0.54	0.25	−0.39	0.57
		ϕ_1	0.90	0.85	−0.047	0.11	0.08	−0.44	0.11
		θ_0	−0.30	−0.10	0.200	0.57	0.26	0.35	0.60
		θ_1	−0.70	−0.68	0.024	0.16	0.11	0.15	0.16
	$\phi_0 = 0.6$	p_{00}	0.99	0.99	−0.003	0.01	0.01	−0.21	0.01
		p_{10}	0.01	0.01	0.001	0.01	0.01	0.14	0.01
		μ_{y0}	3.00	3.01	0.005	0.15	0.12	0.04	0.15
		μ_{y1}	4.00	3.98	−0.021	0.39	0.19	−0.05	0.39
		β_0	0.00	−0.00	−0.004	0.06	0.05	−0.07	0.06
		β_1	0.50	0.50	0.000	0.12	0.08	0.00	0.12
		σ_ζ^2	0.25	0.25	−0.003	0.02	0.02	−0.19	0.02
		ϕ_0	0.60	0.48	−0.122	0.23	0.17	−0.52	0.26
		ϕ_1	0.90	0.84	−0.056	0.14	0.09	−0.41	0.15
		θ_0	−0.30	−0.19	0.108	0.26	0.18	0.42	0.28
		θ_1	−0.70	−0.66	0.038	0.18	0.13	0.21	0.18

6.2. *Simulation Results*

The convergence rates across all the conditions were very high, yielding at most seven non-convergent cases out of 100 replications under each condition. The computational time (namely CPU time) for a single time series was approximately 200 seconds on average (using a Macintosh machine with 2.16 GHz Intel Core 2 Duo CPU). The computational time is expected to increase with greater time series lengths, however.

Summary statistics for all parameter estimates obtained over 100 Monte Carlo replications are shown in Table 5 (for MA coefficient during the deactivation phase, θ_0, of -0.3) and Table 6 (for MA coefficient during the deactivation phase, θ_0, of -0.5). Overall, the parameters were recovered accurately across all the conditions. The root mean squared errors (RMSEs) of most parameters were found to be smaller when the probability of staying in the same regime was smaller (i.e., $p_{00} = p_{11} = 0.90$ as compared with $p_{00} = p_{11} = 0.99$) and when the AR coefficient, ϕ_0, was smaller (i.e., $\phi_0 = 0.3$ as compared with $\phi_0 = 0.6$). When the probability of staying in the same regime was as high as 0.99, relatively few data points were available from the other regime to accurately recover the corresponding parameters. In addition, it is generally harder to obtain accurate parameter estimates when the AR parameter is too close to the nonstationary region (i.e., the AR range that would lead to mean and covariance functions that vary over time, which, for an ARMA(1,1) process, occurs when $|\phi| > 1$). The higher ϕ_0 also led to very similar observed dynamics between the two regimes, thus making the two regimes less separable from each other. Consequently, it is harder to differentiate one regime from the other when $\phi_0 = 0.6$. However, even under the worst situation out of the eight conditions (i.e., $p_{00} = p_{11} = 0.99$ and $\phi_0 = 0.6$), the corresponding biases and RMSEs were not very large. For instance, across all conditions, the largest absolute bias was 0.293 and the largest RMSE was 0.636. Since the true

TABLE 5.
(*Continued*)

p_{00}	ϕ_0	Parameter	$\boldsymbol{\psi}$	$\bar{\hat{\boldsymbol{\psi}}}$	Bias	Emp. SE	Ave. SE	rBias	RMSE
$p_{00} = 0.90$	$\phi_0 = 0.3$	p_{00}	0.90	0.90	0.004	0.02	0.02	0.26	0.02
		p_{10}	0.10	0.10	0.002	0.02	0.02	0.08	0.02
		μ_{y0}	3.00	3.00	0.003	0.09	0.08	0.03	0.09
		μ_{y1}	4.00	4.10	0.096	0.19	0.16	0.51	0.21
		β_0	0.00	0.00	0.001	0.04	0.03	0.02	0.04
		β_1	0.50	0.47	−0.029	0.07	0.06	−0.39	0.08
		σ_ζ^2	0.25	0.26	0.013	0.02	0.02	0.71	0.02
		ϕ_0	0.30	0.39	0.085	0.22	0.15	0.40	0.23
		ϕ_1	0.90	0.89	−0.014	0.07	0.07	−0.19	0.08
		θ_0	−0.30	−0.36	−0.055	0.22	0.16	−0.25	0.23
		θ_1	−0.70	−0.61	0.087	0.12	0.11	0.73	0.15
	$\phi_0 = 0.6$	p_{00}	0.90	0.90	0.005	0.02	0.02	0.28	0.02
		p_{10}	0.10	0.10	0.001	0.02	0.02	0.03	0.02
		μ_{y0}	3.00	3.00	−0.003	0.14	0.13	−0.02	0.14
		μ_{y1}	4.00	4.08	0.077	0.21	0.17	0.36	0.23
		β_0	0.00	0.00	0.002	0.06	0.05	0.03	0.06
		β_1	0.50	0.48	−0.023	0.08	0.07	−0.29	0.08
		σ_ζ^2	0.25	0.27	0.015	0.02	0.02	0.79	0.02
		ϕ_0	0.60	0.60	0.001	0.13	0.11	0.01	0.13
		ϕ_1	0.90	0.87	−0.029	0.10	0.06	−0.28	0.11
		θ_0	−0.30	−0.27	0.031	0.15	0.13	0.20	0.16
		θ_1	−0.70	−0.60	0.104	0.14	0.10	0.77	0.17

$\boldsymbol{\psi}$ = the true values of the parameters.

$\bar{\hat{\boldsymbol{\psi}}} = \frac{1}{N}\sum_{i=1}^{N} \hat{\boldsymbol{\psi}}_i$, the mean of the parameter estimates averaged across 100 replications.

Bias = absolute bias = $\bar{\hat{\boldsymbol{\psi}}} - \boldsymbol{\psi}$.

Emp. SE = empirical standard errors of the estimates across 100 replications.

Ave. SE = average standard errors obtained from observed information matrix across 100 replications.

rBias = relative bias = Bias/SE.

RMSE: = root mean squared error of the estimates = $\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\boldsymbol{\psi}}_i - \boldsymbol{\psi})^2}$.

parameters space in the simulation matched the range of parameter estimates from our empirical example, these results help add credence to the results of our empirical example.

7. Discussion

In the present article, we illustrated the utility of regime-switching state-space models in representing the activation patterns of EMG, as well as the time-varying relationship between EMG and self-report data. This class of modeling tools provides a systematic mechanism to probabilistically detect turning points in a dynamical process, as opposed to resorting to arbitrary cut-off criteria or the opinions of expert coders to accomplish this goal, which may yield inconsistent conclusions across studies.

Results from our empirical application suggested that when differences across regimes are ignored, the regression slope estimates from the baseline (BASE) model can be regarded as a weighted average of the regression slopes from the two regimes postulated in the regime-switching models (i.e., the RS model, the RS-ARMA model, and the final selected model). Since

TABLE 6.
Parameter estimates when $\theta_0 = -0.5$.

p_{00}	ϕ_0	Parameter	$\boldsymbol{\psi}$	$\bar{\hat{\boldsymbol{\psi}}}$	Bias	Emp. SE	Ave. SE	rBias	RMSE
$p_{00} = 0.99$	$\phi_0 = 0.3$	p_{00}	0.99	0.99	−0.003	0.01	0.01	−0.20	0.01
		p_{10}	0.01	0.01	0.001	0.01	0.01	0.13	0.01
		μ_{y0}	3.00	3.01	0.009	0.09	0.06	0.11	0.09
		μ_{y1}	4.00	3.97	−0.033	0.64	0.20	−0.05	0.64
		β_0	0.00	0.00	−0.005	0.04	0.03	−0.12	0.04
		β_1	0.50	0.50	0.005	0.19	0.08	0.03	0.19
		σ_ζ^2	0.25	0.25	−0.003	0.02	0.02	−0.15	0.02
		ϕ_0	0.30	0.22	−0.084	0.38	0.21	−0.22	0.39
		ϕ_1	0.90	0.85	−0.047	0.12	0.08	−0.40	0.13
		θ_0	−0.50	−0.41	0.086	0.41	0.21	0.21	0.42
		θ_1	−0.70	−0.67	0.031	0.16	0.11	0.19	0.16
	$\phi_0 = 0.6$	p_{00}	0.99	0.99	−0.003	0.01	0.01	−0.20	0.01
		p_{10}	0.01	0.01	0.001	0.01	0.01	0.12	0.01
		μ_{y0}	3.00	3.01	0.008	0.13	0.09	0.06	0.13
		μ_{y1}	4.00	3.99	−0.005	0.33	0.18	−0.02	0.33
		β_0	0.00	0.00	−0.004	0.06	0.04	−0.07	0.06
		β_1	0.50	0.50	−0.004	0.11	0.07	−0.04	0.11
		σ_ζ^2	0.25	0.25	−0.003	0.02	0.02	−0.18	0.02
		ϕ_0	0.60	0.31	−0.293	0.44	0.24	−0.67	0.53
		ϕ_1	0.90	0.85	−0.053	0.13	0.08	−0.41	0.14
		θ_0	−0.50	−0.22	0.275	0.46	0.25	0.60	0.54
		θ_1	−0.70	−0.67	0.034	0.17	0.12	0.20	0.17

EMG data segments with bursts are relatively short and are interspersed with long periods of nonactivation, the average time-invariant slope estimates from the baseline model tend to be close to 0 (indicating deactivation). Thus, the brief but important period during which EMG and self-report data show synchronous changes tend to be overlooked. Since facial EMG signals have rarely been examined as a dynamic process, most researchers have not realized that traditional statistical analysis methods, such as simple linear regression, may not be able to capture the time-varying relationship between EMG and other markers of emotions. The need to include regime-dependent AR and MA coefficients further indicated that the changes in facial EMG signals are likely heterogeneous across regimes and over time. In addition, since the model-implied variance of the EMG process under study is also a function of the AR and MA coefficients, such regime-dependent parameters also allow researchers to capture the heterogeneity in the variance of EMG signals across regimes.

We showed that the estimates from the proposed regime-switching models can be used to derive affect-related individual difference characteristics such as rise time and longest burst duration. As indices of EMG temporal features, rise times and longest burst durations were found to be substantially different across participants. Currently, little is known about the factors that govern these individual differences. Future studies can further investigate whether such temporal properties are related to other individual difference characteristics, such as personality and coping strategies.

Results from our simulation study showed that the proposed RS-ARMA model performed well under different conditions. Greater accuracy and precision were found in conditions with lower probabilities of staying within the same regime, and greater between-regime differences in AR coefficients. In other words, regime-switching models tend to perform well when the differences in dynamics across regimes are more pronounced. In terms of standard errors, simulation

TABLE 6.
(*Continued*)

p_{00}	ϕ_0	Parameter	$\boldsymbol{\psi}$	$\bar{\hat{\boldsymbol{\psi}}}$	Bias	Emp. SE	Ave. SE	rBias	RMSE
$p_{00} = 0.90$	$\phi_0 = 0.3$	p_{00}	0.90	0.90	0.004	0.02	0.02	0.27	0.02
		p_{10}	0.10	0.10	0.002	0.02	0.02	0.08	0.02
		μ_{y0}	3.00	3.00	0.001	0.07	0.06	0.02	0.07
		μ_{y1}	4.00	4.09	0.094	0.19	0.16	0.49	0.21
		β_0	0.00	0.00	0.001	0.03	0.03	0.04	0.03
		β_1	0.50	0.47	−0.027	0.08	0.06	−0.36	0.08
		σ_ζ^2	0.25	0.27	0.017	0.02	0.02	0.94	0.03
		ϕ_0	0.30	0.33	0.026	0.21	0.16	0.12	0.21
		ϕ_1	0.90	0.89	−0.012	0.07	0.06	−0.17	0.07
		θ_0	−0.50	−0.47	0.028	0.21	0.15	0.13	0.21
		θ_1	−0.70	−0.60	0.097	0.12	0.11	0.81	0.15
	$\phi_0 = 0.6$	p_{00}	0.90	0.91	0.005	0.02	0.02	0.30	0.02
		p_{10}	0.10	0.10	0.000	0.02	0.02	0.02	0.02
		μ_{y0}	3.00	3.00	−0.004	0.11	0.10	−0.03	0.11
		μ_{y1}	4.00	4.08	0.082	0.20	0.17	0.42	0.21
		β_0	0.00	0.00	0.002	0.05	0.04	0.04	0.05
		β_1	0.50	0.48	−0.024	0.08	0.07	−0.31	0.08
		σ_ζ^2	0.25	0.27	0.020	0.02	0.02	1.03	0.03
		ϕ_0	0.60	0.55	−0.049	0.21	0.13	−0.23	0.22
		ϕ_1	0.90	0.87	−0.026	0.10	0.07	−0.25	0.10
		θ_0	−0.50	−0.40	0.101	0.23	0.15	0.43	0.25
		θ_1	−0.70	−0.59	0.109	0.14	0.11	0.78	0.18

$\boldsymbol{\psi}$ = the true values of the parameters.

$\bar{\hat{\boldsymbol{\psi}}} = \frac{1}{N}\sum_{i=1}^{N} \hat{\boldsymbol{\psi}}_i$, the mean of the parameter estimates averaged across 100 replications.

Bias = absolute bias = $\bar{\hat{\boldsymbol{\psi}}} - \boldsymbol{\psi}$.

Emp. SE = empirical standard errors of the estimates across 100 replications.

Ave. SE = average standard errors obtained from observed information matrix across 100 replications.

rBias = relative bias = Bias/SE.

RMSE: = root mean squared error of the estimates = $\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{\boldsymbol{\psi}}_i - \boldsymbol{\psi})^2}$.

results indicated that a larger number of observations (more than 500 observations) is required to yield more precise estimates of the AR and MA coefficients, especially when the frequency of switching between regimes is extremely low (e.g., $p_{01} = 0.01$). In fact, the transition probabilities of switching between regimes were set to be relatively low (i.e., $p_{01} = 0.01$ or $p_{01} = 0.10$) in the present context to mirror the estimates from the empirical application. In future studies, the performance of the RS-ARMA model with less extreme values of transition probabilities can be further investigated. Our expectation is that the standard errors of the AR and MA coefficients would become smaller as the dependent variable alternates more frequently between regimes. In practice, this suggests that applied researchers may want to consider extending their data collection span to allow for longer recovery periods. Designs in which experimental stimuli are interspersed with segments of recovery period can also increase the probability of switching between regimes. Taken as a whole, researchers are always recommended to plot their time series data and use other exploratory tools (e.g., Hsieh, Ferrer, Chen, & Chow, 2010) to determine whether potential regimes exist prior to model fitting.

The state-space regime-switching model is a very general and flexible modeling framework that includes a variety of linear models as special cases. However, it also has its own limitations. First, as the number of possible regimes increases, there is also a substantial increase in the number of parameters to be estimated. Consequently, the resultant models are very likely unidentified and the number of nonconvergent cases may also increase substantially. Second, compared to standard state-space models with no regime switching, state-space regime-switching models are characterized by increased computational costs. The increase in computational costs is particularly pronounced in cases involving large time series lengths and large numbers of regimes. Third, the likelihood function of a regime-switching model is typically characterized by multiple local maxima. The resulting parameter estimates are, therefore, sensitive to parameter starting values. As a result, it is recommended to use multiple sets of starting values to check if the corresponding estimation results have converged to the same values.

All the models presented in this article were fitted using our own scripts written in MATLAB. Other statistical programs that can handle matrix operations, such as R (R Development Core Team, 2009), SAS/IML (SAS Institute Inc., 2008), GAUSS (Aptech Systems Inc., 2009) and OxMetrics (Timberlake Consultants Ltd., 2009), may also be used. Kim and Nelson (1999a), for instance, provided some GAUSS codes for fitting the models considered in their book. In addition, the BASE model (Model 0) can be fitted using canned routines from several commonly used statistical programs, such as the ARIMA function in R and the ARIMA procedure in SAS. The RS and RS-ARMA models (Models 1 and 2), however, require statistical packages that can handle both Markov-switching models and ARMA components. Mplus (Muthén & Muthén, 2001), Mx (Neale, Boker, Xie, & Maes, 2002) and the msm package in R (Jackson, 2009) can be used to fit hidden Markov models, but structural equation modeling programs are typically not conducive for handling a very large number of observations (e.g., $T = 700$) whereas the msm package has to be adapted specifically to accommodate ARMA processes. On the contrary, the canned ARIMA routines in R and SAS do not allow for regime-switching (or finite mixture) components. In addition, the canned routines from many statistical programs may not have enough flexibility in the syntax structure to readily allow for the inclusion of multiple dependent variables.

In the present study, we chose to perform model fitting at the individual level. One alternative approach is to consider multiple-subject extensions of the regime-switching models by extending the prediction error decomposition function to sum over individuals.[6] This approach is especially useful in cases involving panel data, where insufficient information is available from each individual to warrant model fitting at the individual level. For instance, Dolan, Schmittmann, Lubke, and Neale (2005) presented a regime-switching latent growth curve model and allowed the responses to condition on latent classes. Schmittmann, Dolan, Van der Maas, and Neale (2005) combined the Markovian transition models (which are closely related to regime-switching models) with covariance structure models to examine qualitative changes in cognitive performance over time. These relatively recent developments offer promising alternatives for consolidating longitudinal data with relatively few measurement occasions from multiple subjects. In contrast, constructing a multiple-subject model that is appropriate at the group level for psychophysiological series with a relatively large number of time points (e.g., EMG data) can be tricky due to the

[6] In the context of our EMG data, Participants 1 and 4 were found to show largely similar dynamics. Specifically, the same final selected model was found to characterize their data well. Although the results are omitted here due to space constraints, we did consider a multiple-subject version of the final selected model (i.e., the RS-ARMA model) using pooled data from the two participants. Estimation was done by modifying the log-likelihood function in Equation (B.1) into the form of Equation (B.2) in Appendix B and constraining all the time-invariant parameters to be equal across the two participants. The smoothed probabilities and other latent variables were estimated in the same way as in the single-subject case. This multi-subject model performed well in terms of model fit and yielded group-based parameter estimates that were weighted average values of those obtained from model fitting at the individual level.

disparities in the individuals' change patterns, number of regimes, and frequency of switching between regimes.

Other than the ML method used in the present article, Bayesian methods for handling dynamic models have also become increasingly prevalent over the last decade. Examples utilizing Markov Chain Monte Carlo (MCMC) techniques in fitting state-space models with regime-switching and other related variations can be found, for instance, in the econometric and statistical literature (Kim & Nelson, 1999b; Frühwirth-Schnatter, 2001, 2006). In the psychometric literature, Bayesian methods have been used to fit time series models such as dynamic factor models (Ram et al., 2005; Zhang, Hamaker, & Nesselroade, 2008). Compared to traditional maximum likelihood estimation, MCMC and other simulation-based procedures (Chow, Ferrer, & Hsieh, 2009; Doucet, de Freitas, & Gordon, 2001; West & Harrison, 1997) offer many new possibilities for handling more sophisticated models, including models that are non-Gaussian and/or nonlinear. In future research, it is worthwhile to compare Bayesian methods with traditional ML methods, especially when complex regime switching models with more than two regimes and/or multivariate latent variables are involved.

In this study, univariate time series models were adopted for reasons of parsimony. In particular, EMG data, as opposed to self-report data, were chosen to be the dependent variable. While this choice is somewhat arbitrary, it was motivated by a few practical concerns. For one, EMG data offer more time precision for defining affect-related characteristics such as rise time and the longest burst duration. For another, the participants in the present study were simply unable to attend to the more subtle affective changes that unfolded from seconds to seconds. More often, the participants were only able to detect discrete, more abrupt changes in their negative emotion. As a result, the self-report ratings from the present study were observed to show discrete shifts in a stepwise manner. Vector autoregressive moving average (VARMA; Hamilton, 1994; Lütkepohl, 2005) models that include both EMG and self-reports as dependent variables may not be appropriate in this case due to their inability to capture such discrete shifts. This modeling approach is still useful in other contexts (e.g., involving self-report data that are measured over days), however.

Regime-switching models are not limited in application to the study of emotions. Many psychological processes have been observed to exhibit qualitatively distinct dynamics during different periods. For example, according to Piaget's (1969) theory, human cognitive development is posited to comprise several discrete stages. In the field of cognitive development and learning processes, qualitative shifts are commonly observed (Van der Maas & Molenaar, 1992; Hosenfeld, 1997; Fukuda & Ishihara, 1997; Van Dijk & Van Geert, 2007). While traditional statistical methods (i.e., repeated-measure ANOVA or simple linear regression) usually assume homogeneity of variance over time, regime-switching models provide a flexible way to account for nonnormality and heterogeneity in change patterns. As noted by Cacioppo and Tassinary (1990), one of the obstacles to making inferences using physiological signals stems from the fact that advances in the analysis of complex physiological signals have lagged behind advances in signal acquisition. The regime-switching state-space models proposed in the present study offer a new way of examining existing research questions and in doing so, may help establish new affect-related theories.

Appendix A. The Key Procedures of the Kim Filter

We outline the key procedures for implementing the Kim filter here and refer the readers to Kim and Nelson (1999a) for further details. The Kim filter algorithm can be decomposed into three parts: the Kalman filter, the Hamilton filter and the collapsing process. For didactic reasons, we will describe the Kalman filter followed by the collapsing process and finally, the Hamilton

filter. In actual implementation, however, the Hamilton filter step has to be executed before the collapsing process takes place.

A.1. The Algorithm of the Kalman Filter

The Kalman filter essentially provides a way to derive longitudinal factor or latent variable scores in real time as a new observation, y_t, is brought in. Let $\boldsymbol{\alpha}_{t|t-1}^{i,j} = E(\boldsymbol{\alpha}_t | S_t = j, S_{t-1} = i, \mathbf{Y}_{t-1})$ denote the filtered latent variable estimates based on information up to time $t-1$ and $\mathbf{P}_{t|t-1}^{i,j} = \text{Cov}(\boldsymbol{\alpha}_t | S_t = j, S_{t-1} = i, \mathbf{Y}_{t-1})$ denote the associated covariance matrix. Let $\mathbf{v}_t^{i,j}$ be defined as the one-step-ahead prediction errors and $\mathbf{f}_t^{i,j}$ are their associated covariance matrix. In addition, $\boldsymbol{\alpha}_{t|t}^{i,j}$ and $\mathbf{P}_{t|t}^{i,j}$ can be regarded as the intermediate latent variable estimates and their covariance matrix based on information up to the current observation. The Kalman filter can be expressed as

$$\boldsymbol{\alpha}_{t|t-1}^{i,j} = \mathbf{c}_j + \mathbf{F}_j \boldsymbol{\alpha}_{t-1|t-1}^{i}, \tag{A.1}$$

$$\mathbf{P}_{t|t-1}^{i,j} = \mathbf{F}_j \mathbf{P}_{t-1|t-1}^{i} \mathbf{F}_j' + \mathbf{Q}_j, \tag{A.2}$$

$$\mathbf{v}_t^{i,j} = \mathbf{y}_t - \mathbf{H}_j \boldsymbol{\alpha}_{t|t-1}^{i,j} - \mathbf{d}_j - \mathbf{A}_j \mathbf{x}_t, \tag{A.3}$$

$$\mathbf{f}_t^{i,j} = \mathbf{H}_j \mathbf{P}_{t|t-1}^{i,j} \mathbf{H}_j' + \mathbf{R}_j, \tag{A.4}$$

$$\boldsymbol{\alpha}_{t|t}^{i,j} = \boldsymbol{\alpha}_{t|t-1}^{i,j} + \mathbf{K}_j \mathbf{v}_t^{i,j}, \tag{A.5}$$

$$\mathbf{P}_{t|t}^{i,j} = \mathbf{P}_{t|t-1}^{i,j} - \mathbf{K}_j \mathbf{H}_j \mathbf{P}_{t|t-1}^{i,j}, \tag{A.6}$$

where $\mathbf{K}_j = \mathbf{P}_{t|t-1}^{i,j} \mathbf{H}_j' [\mathbf{f}_t^{i,j}]^{-1}$ is called the Kalman gain; i and j are indices for the previous regime and current regime, respectively. The estimates of $\boldsymbol{\alpha}_{t|t}^{i,j}$ and $\mathbf{P}_{t|t}^{i,j}$ from Equations (A.5) and (A.6) can then be inserted into Equations (A.1) and (A.2) to obtain the prediction results for the next time point. Thus, the filtering procedure works recursively (i.e., latent variable scores are updated sequentially one measurement occasion at a time) from time 1 to T until latent variable estimates, $\boldsymbol{\alpha}_{t|t}^{i,j}$ and $\mathbf{P}_{t|t}^{i,j}$, have been computed over all time points. To start the filter, the initial state $\boldsymbol{\alpha}_0$ with expectation $\boldsymbol{\alpha}_{0|0}$ and covariance matrix $\mathbf{P}_{0|0}$ are needed. Typically, $\boldsymbol{\alpha}_0$ is assumed to have a *diffuse prior* density, which means $\boldsymbol{\alpha}_{0|0}$ is fixed at an arbitrary value and the diagonal elements of the covariance matrix $\mathbf{P}_{0|0}$ are set to some arbitrarily large constant values.

A.2. The Collapsing Process

Given that the total number of regimes is M, note that in the above algorithm, the number of situations to be considered in each iteration will be M times larger than that in the previous iteration. For example, suppose $M = 3$, then $i, j = 1, 2, 3$. At the first iteration, the values of the latent state $\boldsymbol{\alpha}$ are contingent on a total of $3 \times 3 = 9$ values of $\boldsymbol{\alpha}_{t|t-1}^{i,j}$ in Equation (A.1), corresponding to every possible value for i and j. These 9 values of $\boldsymbol{\alpha}_{t|t-1}^{i,j}$ are inserted into Equation (A.5), yielding 9 values of $\boldsymbol{\alpha}_{t|t}^{i,j}$. Subsequently, at the next iteration, $\boldsymbol{\alpha}_{t-1|t-1}^{i}$ in Equation (A.1) is replaced by the 9 values of $\boldsymbol{\alpha}_{t|t}^{i,j}$ obtained from the previous iteration, yielding 3×9 values of $\boldsymbol{\alpha}_{t|t-1}^{i,j}$ in Equation (A.1) at the second iteration. As a result, the number of possible values of filtered estimates $\boldsymbol{\alpha}_{t|t-1}^{i,j}$ will increase in an exponential fashion with respect to time, which will bring insurmountable difficulties both computationally and from the perspective of information storage.

Several approaches have been proposed to circumvent this computational issue. Most of these methods revolve around finding ways to collapse computational terms at the end of each iteration. For instance, the $M \times M$ terms of $\boldsymbol{\alpha}_{t|t}^{i,j}$ and $\mathbf{P}_{t|t}^{i,j}$ from Equations (A.5) and (A.6) can be collapsed into M terms of $\boldsymbol{\alpha}_{t|t}^{j}$ and $\mathbf{P}_{t|t}^{j}$, which can then be inserted into Equations (A.1) and (A.2) as $\boldsymbol{\alpha}_{t-1|t-1}^{i}$ and $\mathbf{P}_{t-1|t-1}^{i}$ to start the next iteration. By doing so, only M terms need to be stored at the end of each iteration. Given $\boldsymbol{\alpha}_{t|t}^{j} = E(\boldsymbol{\alpha}_t|S_t = j, \mathbf{Y}_t)$, this term is treated as a weighted average of $\boldsymbol{\alpha}_{t|t}^{i,j}$ where $i, j = 1, 2, \ldots, M$. It can be shown that

$$\boldsymbol{\alpha}_{t|t}^{j} = \sum_{i=1}^{M} \mathbf{W}_t \boldsymbol{\alpha}_{t|t}^{i,j}, \tag{A.7}$$

$$\mathbf{P}_{t|t}^{j} = \sum_{i=1}^{M} \mathbf{W}_t \left[\mathbf{P}_{t|t}^{i,j} + \left(\boldsymbol{\alpha}_{t|t}^{j} - \boldsymbol{\alpha}_{t|t}^{i,j}\right)\left(\boldsymbol{\alpha}_{t|t}^{j} - \boldsymbol{\alpha}_{t|t}^{i,j}\right)'\right], \tag{A.8}$$

where

$$\mathbf{W}_t = \frac{\Pr[S_{t-1} = i, S_t = j|\mathbf{Y}_t]}{\Pr[S_t = j|\mathbf{Y}_t]}. \tag{A.9}$$

$\mathbf{W}_t$ is called the *weighting factor*. Once $\mathbf{W}_t$ is computed, it is straightforward to obtain the results of $\boldsymbol{\alpha}_{t|t}^{j}$ and $\mathbf{P}_{t|t}^{j}$. Elements in $\mathbf{W}_t$ are computed using the Hamilton filter.

A.3. The Algorithm of the Hamilton Filter

There are two elements in the weighting factor $\mathbf{W}_t$, namely, $\Pr[S_{t-1} = i, S_t = j|\mathbf{Y}_t]$ and $\Pr[S_t = j|\mathbf{Y}_t]$. The collapsing process in Equations (A.7) and (A.8) cannot proceed unless these two elements are calculated at each time point. Consequently, an algorithm named the Hamilton filter (Hamilton, 1989) is used so that elements in $\mathbf{W}_t$ are obtained and the latent variable estimates can be properly weighted.

Similar to the Kalman filter, the Hamilton filter is also a recursive process and it can be expressed as:

$$\Pr[S_{t-1} = i, S_t = j|\mathbf{Y}_{t-1}] = \Pr[S_t = j|S_{t-1} = i] \times \Pr[S_{t-1} = i|\mathbf{Y}_{t-1}], \tag{A.10}$$

$$f(\mathbf{y}_t|\mathbf{Y}_{t-1}) = \sum_{i=1}^{M}\sum_{j=1}^{M} f(\mathbf{y}_t|S_t = j, S_{t-1} = i, \mathbf{Y}_{t-1}) \Pr[S_{t-1} = i, S_t = j|\mathbf{Y}_{t-1}], \tag{A.11}$$

$$\Pr[S_{t-1} = i, S_t = j|\mathbf{Y}_t] = \frac{f(\mathbf{y}_t|S_t = j, S_{t-1} = i, \mathbf{Y}_{t-1}) \Pr[S_{t-1} = i, S_t = j|\mathbf{Y}_{t-1}]}{f(\mathbf{y}_t|\mathbf{Y}_{t-1})}, \tag{A.12}$$

$$\Pr[S_t = j|\mathbf{Y}_t] = \sum_{i=1}^{M} \Pr[S_{t-1} = i, S_t = j|\mathbf{Y}_t], \tag{A.13}$$

where $\Pr[S_t = j|S_{t-1} = i]$ are *transition probabilities* specified as part of the time-invariant parameters in $\boldsymbol{\psi}$ of a model; $f(\mathbf{y}_t|S_t = j, S_{t-1} = i, \mathbf{Y}_{t-1})$ is a likelihood function based on the normal distribution and it is given by the *prediction error decomposition* function in Equation (12). Equations (A.12) and (A.13) enable us to calculate the weighting factor $\mathbf{W}_t$ in Equation (A.9). To start the filter, the initial probabilities of the latent regime, namely, $\Pr[\mathbf{S}_0|\mathbf{Y}_0]$, are needed. These probabilities can be obtained by computing the steady-state or unconditional probabilities

of S_t. For example, in a first-order Markov-switching process which has two possible regimes, the initial probabilities of S_t can be expressed as

$$\Pr[\mathbf{S}_0 = 1|\mathbf{Y}_0] = \frac{1-\mathbf{p}_{22}}{2-p_{11}-\mathbf{p}_{22}} \quad \text{and} \quad \Pr[\mathbf{S}_0 = 2|\mathbf{Y}_0] = \frac{1-p_{11}}{2-p_{11}-\mathbf{p}_{22}},$$

where

$$p_{11} = \Pr[S_t = 1|S_{t-1} = 1] \quad \text{and} \quad \mathbf{p}_{22} = \Pr[S_t = 2|S_{t-1} = 2].$$

A.4. The Algorithm of the Kim Smoother

Given results from the Kim filter, the Kim smoother can then be used to obtain more accurate latent variable estimates and regime probabilities based on all observed information from the entire sample. This smoothing procedure can be summarized in four steps (Kim & Nelson, 1999a):

1. Run the Kalman filter in Equations (A.1) to (A.6) and record $\boldsymbol{\alpha}_{t|t-1}^{i,j}, \mathbf{P}_{t|t-1}^{i,j}, \boldsymbol{\alpha}_{t|t}^{j}, \mathbf{P}_{t|t}^{j}$ for time 1 to T. Run the Hamilton filter in Equations (A.10) to (A.13) and record $\Pr[S_t = j|\mathbf{Y}_t]$ as well as $\Pr[S_t = j|\mathbf{Y}_{t-1}]$.

2. Use backward recursion from $T-1$ to 1 to get the smoothed probability terms $\Pr[\mathbf{S}_{t+1} = k, S_t = j|\mathbf{Y}_T]$ and $\Pr[S_t = j|\mathbf{Y}_T]$ by

$$\Pr[\mathbf{S}_{t+1} = k, S_t = j|\mathbf{Y}_T] = \frac{\Pr[\mathbf{S}_{t+1} = k|\mathbf{Y}_T]\Pr[S_t = j|\mathbf{Y}_t]\Pr[\mathbf{S}_{t+1} = k|S_t = j]}{\Pr[\mathbf{S}_{t+1} = k|\mathbf{Y}_t]}, \tag{A.14}$$

$$\Pr[S_t = j|\mathbf{Y}_T] = \sum_{k=1}^{M} \Pr[\mathbf{S}_{t+1} = k, S_t = j|\mathbf{Y}_T]. \tag{A.15}$$

Meanwhile, use backward recursion to calculate the smoothed $\boldsymbol{\alpha}_t$ and $\mathbf{P}_t$ given that $S_t = j$ and $\mathbf{S}_{t+1} = k$:

$$\boldsymbol{\alpha}_{t|T}^{j,k} = \boldsymbol{\alpha}_{t|t}^{j} + \widetilde{\mathbf{P}}_t^{j,k}\big(\boldsymbol{\alpha}_{t+1|T}^{k} - \boldsymbol{\alpha}_{t+1|t}^{j,k}\big), \tag{A.16}$$

$$\mathbf{P}_{t|T}^{j,k} = \mathbf{P}_{t|t}^{j} + \widetilde{\mathbf{P}}_t^{j,k}\big(\mathbf{P}_{t+1|T}^{k} - \mathbf{P}_{t+1|t}^{j,k}\big)\widetilde{\mathbf{P}}_t^{j,k}, \tag{A.17}$$

where $\widetilde{\mathbf{P}}_t^{j,k} = \mathbf{P}_{t|t}^{j}\mathbf{F}_k'[\mathbf{P}_{t+1|t}^{j,k}]^{-1}$.

3. Use the smoothed probabilities in Step 2 to collapse $M \times M$ terms of $\boldsymbol{\alpha}_{t|T}^{j,k}$ and $\mathbf{P}_{t|T}^{j,k}$ into M terms of $\boldsymbol{\alpha}_{t|T}^{j}$ and $\mathbf{P}_{t|T}^{j}$ by computing the weighted averages:

$$\boldsymbol{\alpha}_{t|T}^{j} = \sum_{k=1}^{M} \frac{\Pr[\mathbf{S}_{t+1} = k, S_t = j|\mathbf{Y}_T]}{\Pr[S_t = j|\mathbf{Y}_T]}\boldsymbol{\alpha}_{t|T}^{j,k}, \tag{A.18}$$

$$\mathbf{P}_{t|T}^{j} = \sum_{k=1}^{M} \frac{\Pr[\mathbf{S}_{t+1} = k, S_t = j|\mathbf{Y}_T]}{\Pr[S_t = j|\mathbf{Y}_T]}\big[\mathbf{P}_{t|T}^{j,k} + \big(\boldsymbol{\alpha}_{t|T}^{j} - \boldsymbol{\alpha}_{t|T}^{j,k}\big)\big(\boldsymbol{\alpha}_{t|T}^{j} - \boldsymbol{\alpha}_{t|T}^{j,k}\big)'\big]. \tag{A.19}$$

4. Take the weighted average of $\boldsymbol{\alpha}_{t|T}^{j}$ from Step 3 by summing across the M regimes in effect:

$$\boldsymbol{\alpha}_{t|T} = \sum_{j=1}^{M} \Pr[S_t = j|\mathbf{Y}_T]\boldsymbol{\alpha}_{t|T}^{j}. \tag{A.20}$$

These four steps yield two sets of final products: $\boldsymbol{\alpha}_{t|T}$, the smoothed latent variable estimates conditional on all manifest observations, and $\Pr[S_t = j|\mathbf{Y}_T]$, the smoothed probabilities of being in the jth regime conditional on all manifest observations.

Appendix B. The Likelihood Function for a Simple Multiple-Subject Model

At the individual level, the log-likelihood function of the kth participant can be expressed as

$$LL_k = \log\big[f(\mathbf{y}_{k1}, \mathbf{y}_{k2}, \ldots, \mathbf{y}_{kT}; \boldsymbol{\psi}_k)\big] = \sum_{t=1}^{T} \log\big[f(\mathbf{y}_{kt}|\mathbf{Y}_{k,t-1}; \boldsymbol{\psi}_k)\big], \tag{B.1}$$

where $f(\mathbf{y}_{kt}|\mathbf{Y}_{k,t-1}; \boldsymbol{\psi}_k)$ can be obtained as by-products from the Kim filter, and $\boldsymbol{\psi}_k$ represents a vector of time-invariant parameters for the kth participant. Suppose time series observations from different participants are independent of each other, then a multiple-subject extension of the regime-switching models can be structured by taking the sum of log-likelihood functions over individuals. For simplicity, all the time-invariant parameters are constrained to be equal across participants, namely, $\boldsymbol{\psi}_1 = \boldsymbol{\psi}_2 = \cdots = \boldsymbol{\psi}_k = \boldsymbol{\psi}$. Thus, the log-likelihood function of a multiple-subject model with a total of K participants can be expressed as

$$LL = \sum_{k=1}^{K} LL_k = \sum_{k=1}^{K} \sum_{t=1}^{T} \log\big[f(\mathbf{y}_{kt}|\mathbf{Y}_{k,t-1}; \boldsymbol{\psi})\big]. \tag{B.2}$$

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B.N. Petrov, & F. Csaki (Eds.), *Proceedings of the second international symposium on information theory* (pp. 267–281). Budapest: Akademiai Kiado.

Akima, H. (1970a). A method of univariate interpolation that has the accuracy of a third-degree polynomial. *ACM Transactions on Mathematical Software*, *17*(3), 341–366.

Akima, H. (1970b). A new method of interpolation and smooth curve fitting based on local procedures. *Journal of the Association of Computing Machinery*, *17*(4), 589–602.

Aptech Systems Inc. (2009). *GAUSS (Version 10)*. Black Diamond: Author. Computer software manual.

Biopac Systems, Inc. (2005). *MP system hardware guide*, Goleta: Biopac Systems, Inc. Computer software manual.

Cacioppo, J., Martzke, J.S., Petty, R.E., & Tassinary, L.G. (1988). Specific forms of facial EMG response index emotions during an interview: From Darwin to the continuous flow hypothesis of affect-laden information processing. *Journal of Personality and Social Psychology*, *54*, 592–604.

Cacioppo, J., & Petty, R. (1981). Electromyograms as measures of extent and affectivity of information processing. *American Psychologist*, *36*, 441–456.

Cacioppo, J., Petty, R., Losch, M., & Kim, H. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology*, *50*, 260–268.

Cacioppo, J., & Tassinary, L. (1990). Inferring psychological significance from physiological signals. *American Psychologist*, *45*, 16–28.

Chow, S.-M., Ferrer, E., & Hsieh, F. (Eds.) (2009). *Notre Dame series on quantitative methodology: Vol. 4. Statistical methods for modeling human dynamics: An interdisciplinary dialogue*. New York: Taylor & Francis.

Davidson, R. (1998). Affective style and affective disorders: Perspectives from affective neuroscience. *Cognition and Emotion*, *12*(3), 307–330.

Davis, W.J., Rahman, M.A., Smith, L.J., Burns, A., Senecal, L., McArthur, D. et al. (1995). Properties of human affect induced by static color slides (IAPS): Dimensional, categorical and electromyographic analysis. *Biological Psychology*, *41*, 229–253.

Dimberg, U. (1990). Facial electromyography and emotional reactions. *Psychophysiology*, *27*, 481–494.

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*(1), 86–89.

Dolan, C., Schmittmann, V., Lubke, G., & Neale, M. (2005). Regime switching in the latent growth curve mixture model. *Structural Equation Modeling*, *12*(1), 94–119.

Doucet, A., de Freitas, N., & Gordon, N. (Eds.) (2001). *Sequential Monte Carlo methods in practice*. New York: Springer.

Durbin, J., & Koopman, S. (2001). *Time series analysis by state space methods*. Oxford: Oxford University Press.

Fridlund, A., & Cacioppo, J. (1986). Guidlines for human electromyographic research. *Psychophysiology*, *23*(5), 567–589.
Frühwirth-Schnatter, S. (2001). Fully Bayesian analysis of switching Gaussian state space models. *Annals of the Institute of Statistical Mathematics*, *53*(1), 31–49.
Frühwirth-Schnatter, S. (2006). *Finite mixture and Markov switching models*. New York: Springer.
Fukuda, K., & Ishihara, K. (1997). Development of human sleep and wakefulness rhythm during the first six months of life: Discontinuous changes at the 7th and 12th week after birth. *Biological Rhythm Research*, *28*, 94–103.
Gilbert, D. (2007). *Stumbling on happiness*. New York: Vintage.
Gilboa, E., & Revelle, W. (1994). Personality and the structure of affective responses. In S. Van Goozen, N. Van de Poll, & J. Sergeant (Eds.), *Emotions: Essays on emotion theory*. Philadelphia: Lawrence Erlbaum Associates.
Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*, 357–384.
Hamilton, J.D. (1994). *Time series analysis*. Princeton: Princeton University Press.
Harvey, A.C. (1981). *The econometric analysis of time series*. New York: Wiley.
Harvey, A.C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge: Cambridge University Press.
Hodges, P., & Bui, B. (1996). A comparison of computer-based methods for the determination of onset of muscle contraction using electromyography. *Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control*, *101*, 511–519.
Hosenfeld, B. (1997). Indicators of discontinuous change in the development of analogical reasoning. *Journal of Experimental Child Psychology*, *64*, 367–395.
Hsieh, F., Ferrer, E., Chen, S., & Chow, S.-M. (2010). Exploring the dynamics of dyadic interactions via hierarchical segmentation. *Psychometrika*, *75*(2), 351–372.
Jackson, C. (2009). *Package 'msm' (Version 0.9.5)*. Vienna: R Foundation for Statistical Computing. Computer software manual. Available from http://cran.r-project.org/web/packages/msm/index.html.
Johnson, T., Elashoff, R., & Harkema, S. (2003). A Bayesian change-point analysis of electromyographic data: Detecting muscle activation patterns and associated applications. *Biostatistics*, *4*(1), 143–164.
Kim, C.-J., & Nelson, C.R. (1999a). *State-space models with regime switching: Classical and Gibbs-sampling approaches with applications*. Cambridge: MIT Press.
Kim, C.-J., & Nelson, C.R. (1999b). Has the US economy become more stable? A Bayesian approach on a Markov-switching model of the business cycle. *The Review of Econometrics and Statistics*, *81*, 608–616.
Lang, P., Bradley, M., & Cuthbert, B. (2005). *International affective picture system (IAPS): Instruction manual and affective ratings* (Technical Report A–6). The Center for Research in Psychophysiology, University of Florida.
Larsen, R.J. (2000). Toward a science of mood regulation. *Psychological Inquiry*, *11*(3), 129–141.
Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. New York: Springer.
Muthén, L.K., & Muthén, B.O. (2001). *Mplus (Version 5)*. Los Angeles: Muthén and Muthén. Computer software manual.
Neale, M.C., Boker, S.M., Xie, G., & Maes, H.H. (2002). *Mx: Statistical modeling (Version 5)*. Richmond: Virginia Commonwealth University. Computer software manual.
Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.
R Development Core Team. (2009). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Computer software manual. Available from http://www.R-project.org (ISBN 3-900051-07-0).
Ram, N., Chow, S.-M., Bowles, R.P., Wang, L., Grimm, K.J., Fujita, F. et al. (2005). Recovering cyclicity in pleasant and unpleasant affect using spectral analysis, the rating scale model, and planned incompleteness. *Psychometrika: Application Reviews and Case Studies*, *70*, 773–790.
Ravaja, N., Saari, T., Puttonen, S., & Keltikangas Jarvinen, L. (2008). The psychophysiology of James Bond: Phasic emotional responses to violent video game events. *Emotion*, *8*(1), 114–120.
Robinson, J.D., Cinciripini, P.M., Carter, B.L., Lam, C.Y., & Wetter, D.W. (2007). Facial EMG as an index of affective response to nicotine. *Experimental and Clinical Psychopharmacology*, *15*(4), 390–399.
SAS Institute Inc. (2008). *SAS 9.2 help and documentation*. Cary: SAS Institute Inc. Computer software manual.
Schmittmann, V., Dolan, C., Van der Maas, H., & Neale, M. (2005). Discrete latent Markov models for normally distributed response data. *Multivariate Behavioral Research*, *40*(4), 461–488.
Schwartz, G.E. (1975). Biofeedback self-regulation, and the patterning of physiological processes. *American Scientist*, *63*, 314–324.
Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464.
Sison, C., Alpert, M., Fudge, R., & Stern, R. (1996). Constricted expressiveness and psychophysiological reactivity in schizophrenia. *The Journal of Nervous and Mental disease*, *184*(10), 589–597.
Staude, G., & Wolf, W. (1999). Objective motor response onset detection in surface myoelectric signals. *Medical Engineering and Physics*, *21*, 449–467.
The MathWorks, Inc. (1991). *MATLAB user's guide*. Natick: The MathWorks, Inc. Computer software manual.
Timberlake Consultants Ltd. (2009). OxMetrics (Version 6.0). London: Author. Computer software manual.
Van der Maas, H., & Molenaar, P. (1992). Stagewise cognitive development: An application of catastrophe theory. *Psychological Review*, *99*(3), 395–417.
Van Dijk, M., & Van Geert, P. (2007). Wobbles, humps and sudden jumps: A case study of continuity, discontinuity and variability in early language development. *Infant and Child Development*, *16*(1), 7–33.

West, M., & Harrison, J. (1997). *Bayesian forecasting and dynamic models* (2nd ed.). New York: Springer.
Weyers, P., Hlberger, A., Hefele, C., & Pauli, P. (2006). Electromyographic responses to static and dynamic avatar emotional facial expressions. *Psychophysiology*, *43*, 450–453.
Zhang, Z., Hamaker, E.L., & Nesselroade, J.R. (2008). Comparisons of four methods for estimating a dynamic factor model. *Structural Equation Modeling*, *15*(1), 377–402.

Manuscript Received: 2 JUL 2009
Final Version Received: 11 FEB 2010
Published Online Date: 5 AUG 2010