PSYCHOMETRIKA—VOL. 74, NO. 4, 727–745
DECEMBER 2009
DOI: 10.1007/S11336-009-9113-4

USING THRESHOLD AUTOREGRESSIVE MODELS TO STUDY DYADIC INTERACTIONS

ELLEN L. HAMAKER

METHODOLOGY AND STATISTICS, FACULTY OF SOCIAL SCIENCES, UTRECHT UNIVERSITY

ZHIYONG ZHANG

QUANTITATIVE PSYCHOLOGY, UNIVERSITY OF NOTRE DAME

HAN L.J. VAN DER MAAS

PSYCHOLOGICAL METHODOLOGY, UNIVERSITY OF AMSTERDAM

Considering a dyad as a dynamic system whose current state depends on its past state has allowed researchers to investigate whether and how partners influence each other. Some researchers have also focused on how differences between dyads in their interaction patterns are related to other differences between them. A promising approach in this area is the model that was proposed by Gottman and Murray, which is based on nonlinear coupled difference equations. In this paper, it is shown that their model is a special case of the threshold autoregressive (TAR) model. As a consequence, we can make use of existing knowledge about TAR models with respect to parameter estimation, model alternatives and model selection. We propose a new estimation procedure and perform a simulation study to compare it to the estimation procedure developed by Gottman and Murray. In addition, we include an empirical example based on interaction data of three dyads.

Key words: dynamic system, TAR model, autoregressive, dyadic interaction.

The notion of a dynamic system has become increasingly popular in psychology, both as a metaphor and as a quantitative approach. Recently, there has been a burst of applications, for instance in the study of development (Thelen & Smith, 1994; Van der Maas & Molenaar, 1992; Van Geert & Van Dijk, 2002), self-regulation of behavior (Carver & Scheier, 1998), personality (Shoda, Tiernan, & Mischel, 2002), addiction (Warren, Hawkins, & Sprott, 2003; Witkiewitz, Van der Maas, Hufford, & Marlatt, 2007), grief (Bisconti, Bergeman, & Boker, 2004), psychopathology (Granic & Hollenstein, 2003), and psychotherapy (Schiepek, 2003). In addition, several leading journals have devoted special issues to the topic (Vallacher & Nowak, 1997; Vallacher, Read, & Nowak, 2002), and in 1997 a new journal called Nonlinear Dynamics, Psychology, and Life Sciences was started (Guastello, 1997).

Mathematically, a dynamic (or dynamical) system is a set of equations that expresses how the state of a system (represented by one or more variables) changes as a function of its previous state. These equations may be deterministic or stochastic (Clayton, 1997), and they can be defined in continuous time as differential equations, or in discrete time as difference equations. A particularly useful application of dynamic system theory in psychological research concerns the study of dyadic interactions: Representing each partner by one or more variables and measuring them repeatedly over time allows the researcher to determine *whether* and *how* the partners influence each other.

This study was supported by the National Institute on Aging (Grant 5T32AG020500), and the Dutch Organization for Scientific Research (NWO; VENI Grant 451-05-012).

Requests for reprints should be sent to Ellen L. Hamaker, Methods and Statistics, Faculty of Social Sciences, Utrecht University, P.O. Box 80140, 3508 TC, Utrecht, The Netherlands. E-mail: e.l.hamaker@uu.nl

© 2009 The Psychometric Society. This article is published with open access at Springerlink.com

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

Several researchers have emphasized the need for nonlinear models to study complex systems such as dyads (Chow, Ferrer, & Nesselroade, 2007; Gottman, Murray, Swanson, Tyson, & Swanson, 2002; Van der Maas & Raijmakers, 2000; Olthof, Kunnen, & Boom, 2000). A promising nonlinear approach to dyadic interactions is the model proposed by Gottman and Murray (GM model hereafter; Cook, Tyson, White, Gottman, & Murray 1995). This approach demonstrated its merits when two studies showed that parameters obtained with the GM model from a 15 minute dyadic interaction were predictive of marital status and marital happiness three to six years later (Cook et al., 1995; Gottman, Swanson, & Murray, 1999).

Apart from its usefulness in predicting divorce, the GM model has formed the basis for a new form of couples therapy referred to as the Gottman Method Couples Therapy (Gottman, Coan, Carrere, & Swanson, 1998; Schwartz Gottman, 2004). Moreover, research based on the GM model has uncovered differences in the interaction patterns of homosexual dyads versus heterosexual dyads, as well as differences between homosexual males and homosexual females (Gottman, Levenson, Swanson, Swanson, Tyson, & Yoshimoto, 2003), which may have important implications for effective couple's therapy for homosexual couples. Gottman et al. (2002) extended the GM model to triadic interactions for parents and their baby, which makes the model suitable to study the transition of newlyweds to parenthood. In addition, the GM model has been shown to work not only with behavioral data, but also with self-rating data and physiological data, as well as combinations thereof (Gottman et al., 2002).

Although the emphasis in GM modeling has been on dyadic interactions, the applicability of the GM model is in no way restricted to dyadic or observational data. That is, the GM model could just as easily be used to investigate other dynamic systems in psychology, ranging from emotion regulatory systems, to the reciprocal influence of physical aspects such as heart rate and breathing.

However, there are some unresolved issues regarding parameter estimation and model selection that may forestall this model from becoming a popular dynamic system approach in psychological research. Therefore, the aim of the current paper is to solve what we consider the most urgent issues. In the following section, we present the threshold autoregressive (TAR) model, which is a nonlinear time series model. This is followed by a section in which we discuss the GM model, its interpretation and use. In the third section, we show that the GM model is a special case of the TAR model, and we formulate a number of TAR-based alternatives for the GM model that may be useful in the study of dyadic interactions. In the fourth section, we discuss the existing estimation procedure for the GM model, and suggest a new estimation procedure, which we compare in the fifth section using a simulation study. The sixth section consists of an empirical study in which we show how to select the best model from several alternatives. We end this paper with a discussion of the GM model in across-dyad comparisons.

1. Tong's Threshold Autoregressive Models

The TAR model, introduced by Tong and Lim (1980), consists of two or more (linear) AR processes, which we refer to as regimes, and a switching mechanism that determines in which of these regimes the system is at each occasion. This switching mechanism depends on the value taken on by a manifest threshold variable. While the TAR model is nonlinear in time, it can be thought of as a piece-wise linearization in state-space.

Let y_t denote a TAR process, and let z_t denote the threshold variable such that the value that z_{t-d} takes on determines in which regime y_t is. The parameter d is referred to as the delay. Let $\{R^{(m)}\}$ form a nonoverlapping partition of $(-\infty, \infty)$, that is, $R^{(m)} = (\tau^{(m-1)}, \tau^{(m)}]$ with $-\infty = \tau^{(0)} < \tau^{(1)} < \cdots < \tau^{(M)} = \infty$ (Fan & Yao, 2003), meaning there are M different regimes, and $M - 1$ thresholds of interest, that is, $\tau^{(1)}$ to $\tau^{(M-1)}$. The variable y_t is regressed on itself

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

at previous occasions and on some exogenous variable x_t at this occasion, and earlier occasions. The regression coefficients, the residual variance, and the number of predictors (i.e., previous values of y_t and x_t), may all depend on the regime the system is in. Let $I(z_{t-d} \in R^{(m)})$ denote the indicator function, which equals 1 if $z_{t-d} \in R^{(m)}$, and is 0 otherwise. Then a general expression of the TAR model is

$$y_t = \sum_{m=1}^{M} \left\{ \alpha_0^{(m)} + \sum_{p=1}^{P^{(m)}} \alpha_p^{(m)} y_{t-p} + \sum_{q=0}^{Q^{(m)}} \gamma_q^{(m)} x_{t-q} + \sigma^{(m)} \epsilon_t \right\} I\left(z_{t-d} \in R^{(m)}\right), \tag{1}$$

where ϵ_t is normally distributed with mean 0 and variance 1, $P^{(m)}$ indicates the maximum number of previous observations y_{t-p} and $Q^{(m)}$ indicates the maximum number of previous observations x_{t-q} that are used to predict y_t in regime m.

Tong and Lim (1980) propose three options for the threshold variable z_t. First, one can use the outcome variable (i.e., $y_t = z_t$), resulting in a self-exciting TAR model. This is a model in which there is a feedback loop that governs the behavior of the system. It has been used in psychological research by Warren (2002) and Warren et al. (2003) to model periods of relapse and recovery in alcohol addicts and sex offenders. Second, one can use a different variable as the threshold variable (possibly x_t), resulting in an open loop TAR system. A possible application of this model could consist of considering day until menses as the threshold variable for mood. Third, one can use two variables as each other's threshold variable, such that x_t is the threshold variable for y_t, and vice versa. Such bivariate systems are referred to as a closed loop TAR systems (TARSC; Tong & Lim, 1980), and have been used to model predator-prey cycles. In psychological research, TARSCs can be used to model the mutual effects of two variables that make up a dynamic system together, for instance the two partners of a dyad.

2. Gottman and Murray's Approach to Study Marital Interaction

To study the dyadic interaction of married couples, Gottman and Murray developed the following procedure (Gottman et al., 2002). A couple is asked to have a conversation on a topic that creates a conflict in their relationship. During this conversation they are videotaped and their verbal and nonverbal behavior is scored according to the Specific Affect Coding System (SPAFF; Gottman, McCoy, Coan, & Collier, 1996). At each second both partners are scored, and the scores are then summed over 6 second intervals, resulting in a score for each partner on a scale from −24 (implying extremely negative affect behavior), to 24 (implying extremely positive affect behavior). Typically, a 15-minute conversation is used, resulting in a bivariate series of 150 occasions (Gottman et al., 2002).

2.1. The GM Model

To model the interaction between wife and husband, Gottman and Murray proposed a pair of coupled deterministic equations, one for the wife and one for the husband. Since these two equations are essentially the same with just the partners' roles reversed, only the equations for the wife are presented here.[1] The basic idea is that the wife's current behavior W_t is a function of her behavior at the previous occasion W_{t-1} and her husband's behavior at the previous occasion H_{t-1}, such that

$$W_t = \alpha_{0w} + \alpha_{1w} W_{t-1} + I_W(H_{t-1}), \tag{2}$$

[1]The current equations are based on Gottman et al. (1999). Other presentations of the GM model contain a certain asymmetry because it is based on turn of speech coding, rather than coding per time interval (e.g., Cook et al., 1995).

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

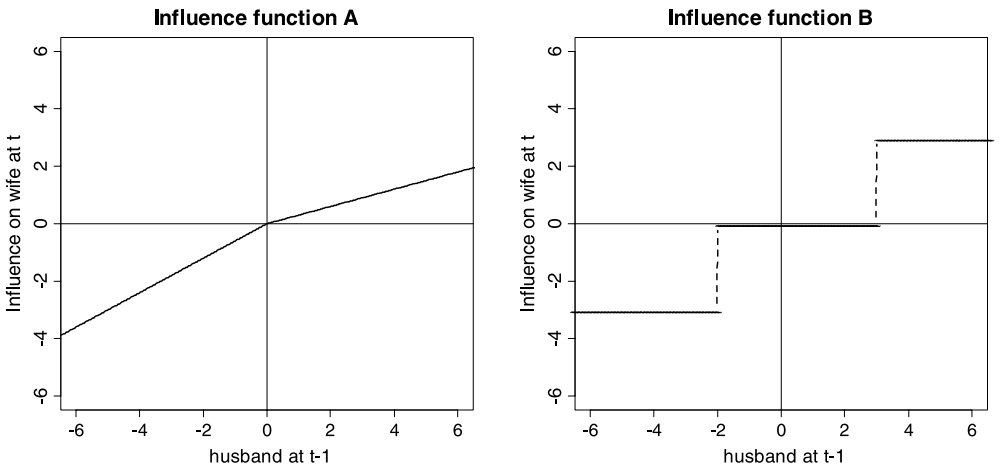


FIGURE 1.
Functions representing the influence of the husband on his wife. The left shows the influence function A with $\gamma_w^- = 0.6$ as the slope of the negative part, and $\gamma_w^+ = 0.3$ as the slope of the positive part. On the right influence function B is depicted, with negative threshold $\tau_w^- = -2$, negative level $\delta_w^- = -3$, positive threshold $\tau_w^+ = 3$, and positive level $\delta_w^+ = 3$.

where α_{0w} is a constant, and α_{1w} is a regression parameter by which the current wife's behavior can be predicted from her behavior at the previous occasion. It is assumed that $0 < \alpha_{1w} < 1$, and this parameter is referred to as the wife's inertia, because the closer this parameter is to 1, the more likely the wife is to maintain her behavior.

2.2. *Influence Functions*

The last term in (2), $I_W(H_{t-1})$, represents the influence of the husband on his wife. Gottman and Murray proposed two alternatives for this influence, which they coined the bilinear influence function and the ojive (or ogive) influence function. However, to avoid confusion with the unrelated bilinear function known in time series literature (cf. Granger & Andersen, 1978), we will refer to these influence functions as A and B in this paper. Both influence functions are illustrated in Figure 1.

Influence function A is based on the idea that more extreme behavior of the husband has a stronger influence on his wife. However, the strength of this influence depends on whether he is positive or negative. For instance, the wife's behavior may be more strongly affected by a particular increase in the husband's behavior when he is being negative than when he is being positive. Hence, we distinguish between a negative regime and a positive regime and write

$$I_W(H_{t-1}) = \begin{cases} \gamma_w^- H_{t-1} & \text{if } H_{t-1} \leq 0, \\ \gamma_w^+ H_{t-1} & \text{if } H_{t-1} > 0, \end{cases} \tag{3}$$

where the superscripts $^-$ and $^+$ refer to the negative and positive regime, respectively.

Influence function B is based on the assumption that influence of the behavior of the husband on his wife is constant until his behavior passes some threshold. Based on the descriptions and plots of this model in Gottman et al. (2002) and Murray (2002), we can express this model as

$$I_W(H_{t-1}) = \begin{cases} \delta_w^- & \text{if } H_{t-1} \leq \tau_w^-, \\ 0 & \text{if } \tau_w^- < H_{t-1} \leq \tau_w^+, \\ \delta_w^+ & \text{if } \tau_w^+ < H_{t-1}, \end{cases} \tag{4}$$

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

where τ_w^- and τ_w^+ are the lower and upper thresholds, respectively. Thus, the influence function is a step function consisting of three regimes: a negative regime (when $H_{t-1} \leq \tau_w^-$), a neutral regime (when $\tau_w^- < H_{t-1} \leq \tau_w^+$), and a positive regime (when $\tau_w^+ < H_{t-1}$). If the husband is behaving (fairly) neutral, his wife is not affected by him, but when his behavior becomes more positive or negative and passes one of the thresholds, it leads to a sudden change in the wife's behavior (i.e., δ_w^- or δ_w^+).

2.3. Long-Term Behavior in the GM Model

Although Gottman et al. (2002) present the GM model as deterministic, they acknowledge that in practice the behavior of a dyad is not deterministic. Hence, the model in (2) should be extended to include a residual term $\epsilon_{w,t}$, for which we may assume a normal distribution with mean zero and variance σ_w^2. This residual can be thought of as a random shock that is given to the system at occasion t, and which continues to affect successive behavior. For instance, the wife could have some association or thought at occasion t that influences her affective behavior at this occasion, but also at following occasions through the regression of W_t on W_{t-1}.

To determine the long-term behavior associated with the GM model, we need to determine the nullclines of the wife and the husband first, that is, the functions that identify the values for which there is no (expected) change in the wife's or the husband's behavior, respectively. Gottman et al. (2002) have derived the nullclines for the deterministic model presented in (2) (see Gottman et al., 2002, p. 136). The derivation of the nullclines for a model with random shocks is given in the Appendix: It shows that these nullclines are the same as for the deterministic GM model presented by Gottman et al. (2002), and are linear transformations of the influence functions.

Nullclines can be thought of as functions in state-space, which is a two-dimensional space with W and H as its axes. The current state of the system can be represented by a point in this two dimensional space with coordinates $\{W_t, H_t\}$. Then the behavior of a system can be described as the motion of this point in the state-space. Each of the nullclines divides the state-space into two areas, which are characterized by different (expected) changes from one occasion to the next. In Figure 2, an example of this is given. The nullcline of the wife (based on influence function B) is depicted on the left and divides the state-space into two areas: an upper area that is characterized by a movement downwards, and a lower area that is characterized by a movement upwards. The length of the arrows indicates the amount of (expected) change from one occasion to the next. Note that at the wife's nullcline itself there is no (expected) movement in the vertical direction, meaning the wife is in equilibrium. Similarly, the nullcline of the husband, depicted in the middle plot of Figure 2, divides the state-space in an area on the left that is characterized by a movement to the right, and an area on the right that is characterized by a movement to the left. At the husband's nullcline itself, there is no (expected) movement in the horizontal direction.

When the nullclines of the husband and the wife intersect, this identifies a point in which both the husband and the wife are in equilibrium, which means the entire system is in equilibrium. These points are referred to as *influenced steady states*, and can be further distinguished into *stable steady states*, which attract the system (meaning that in the neighborhood of this equilibrium the system moves towards it), and *unstable steady states* that repel the system when it gets close to them. This becomes clear from the plot on the right of Figure 2. In this plot, the movements in the vertical direction (i.e., the wife) and in the horizontal direction (i.e., the husband) are combined. It shows that the three intersection of the solid lines form stable steady states, while the intersections of the dashed lines are unstable steady states.

The intersection of the wife's nullcline with the y-axis is referred to as the uninfluenced steady state of the wife because it identifies the wife's equilibrium when her husband does not influence her (i.e., $I_W(H_{t-1}) = 0$). Similarly, the intersection of the husband's nullcline with the x-axis is the husband's uninfluenced steady state. Influenced as well as uninfluenced steady states have been used by Gottman and his colleagues to predict marital status and marital satisfaction.

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

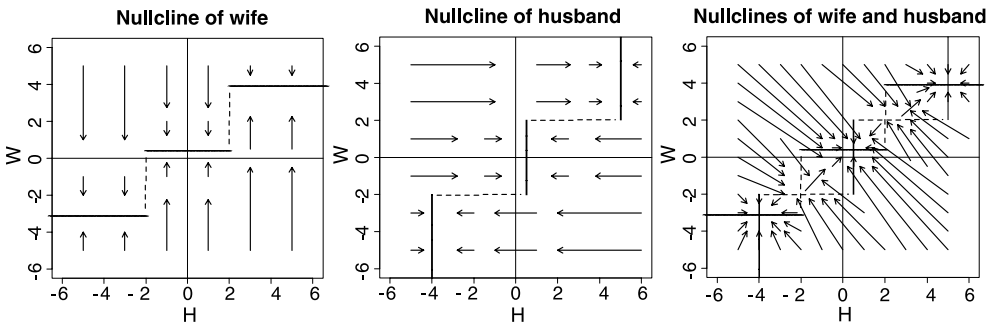


FIGURE 2.
State-space for influence function B. The arrows indicate the change over one time unit. The plot on the left shows the nullcline of the wife, and the arrows show the change in the wife's scores over one time unit. The middle plot includes the nullcline of the husband, and the arrows show the change of the husband's score over time. Combining these two nullclines gives the plot on the right, and the arrows show the change of the system in the state-space over a single time interval. It can be seen that there are five intersections that identify the equilibria of the system, three of which are stable steady states that attract the system, and two of which are unstable steady states that repel the system.

3. GM Model as a TARSC and Some Useful Alternatives

It can be easily shown that the stochastic version of the GM model is a special case of the general TARSC presented in (1). Let y_t be the wife's score, such that we can replace it by W_t in (1), and let $x_t = z_t$ be the husband's score such that it can be replaced by H_t. For both versions of the GM model, we let $P^{(m)} = 1$ for all m (such that only the wife's most recent behavior predicts her current behavior), and $d = 1$ (such that regime-switching depends on the observation at the previous occasion). In the presentation of the two versions of the GM model as TARSCs, we add a subscript w to the parameters to indicate that these belong to the model of the wife (and may differ from those of the husband).

For influence function A (see (3)) we let $M = 2$, such that there are two regimes. In addition, we let $Q^{(m)} = 1$ with $\gamma_{0w}^{(m)} = 0$ for all m (i.e., only H_{t-1} is used in the prediction of W_t). Because α_{0w} and α_{1w} are fixed across both regimes, their regime index m is dropped. This gives us

$$W_t = \alpha_{0w} + \alpha_{1w} W_{t-1} + \sum_{m=1}^{2} \left\{ \gamma_{1w}^{(m)} H_{t-1} + \epsilon_t \right\} I\left(H_{t-1} \in R_w^{(m)}\right), \tag{5}$$

where $R^{(m=1)} = (-\infty, 0]$ and $R^{(m=2)} = (0, \infty]$.

To obtain the GM model with influence function B (see (4)), we set $M = 3$. By fixing $Q^{(m)}$ to zero and keeping α_{1w} fixed across regimes, we get

$$W_t = \alpha_{1w} W_{t-1} + \sum_{m=1}^{3} \left\{ \alpha_{0w}^{(m)} + \epsilon_t \right\} I\left(H_{t-1} \in R_w^{(m)}\right), \tag{6}$$

where $R_w^{(m=1)} = (-\infty, \tau_w^-]$, $R_w^{(m=2)} = (\tau_w^-, \tau_w^+]$, and $R_w^{(m=3)} = (\tau_w^+, \infty]$. The wife's constant α_{0w} from (2) is now equal to the constant of the middle regime (i.e., $\alpha_{0w}^{(2)}$). Other constants in (6) are the sum of the constant in (2) and the constant associated with either the negative or positive regime in (4), that is, $\alpha_{0w}^{(1)} = \alpha_{0w} + \delta_w^-$ and $\alpha_{0w}^{(3)} = \alpha_{0w} + \delta_w^+$.

Based on the general expression in (1), we can formulate many alternatives to the two versions of the GM model. Here we briefly touch upon a few alternatives, which could be of interest for the study of dyadic interactions. First, a slightly more flexible version of the GM

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

model with influence function A includes a freely estimated threshold, rather than fixing it to zero, that is, $R_w^{(m=1)} = (-\infty, 0]$ and $R_w^{(m=2)} = (0, \infty]$ is replaced by $R_w^{(m=1)} = (-\infty, \tau_w]$ and $R_w^{(m=2)} = (\tau_w, \infty]$ such that the point at which the effect of the husband's behavior on his wife changes is not necessarily at his zero score.

Second, we could combine the two influence functions such that there is a step function (as in Influence function B), but within each regime there is also an influence of the partner (as in Influence function A), that is

$$W_t = \alpha_{1w} W_{t-1} + \sum_{m=1}^{3} \left\{ \alpha_{0w}^{(m)} + \gamma_{1w}^{(m)} H_{t-1} + \epsilon_t \right\} I\left(H_{t-1} \in R_w^{(m)}\right), \tag{7}$$

with $R_w^{(m=1)} = (-\infty, \tau_w^-]$, $R_w^{(m=2)} = (\tau_w^-, \tau_w^+]$, and $R_w^{(m=3)} = (\tau_w^+, \infty]$. This would imply that while there are certain thresholds at which there is a sudden change in the amount of influence, there may also be a more gradual effect of the husband's behavior within a particular regime.

Third, the inertia parameter α_{1w} could be allowed to differ across regimes, so that the tendency of a person to remain in a certain state could differ, that is,

$$W_t = \sum_{m=1}^{3} \left\{ \alpha_{0w}^{(m)} + \alpha_{1w}^{(m)} W_{t-1} + \epsilon_t \right\} I\left(H_{t-1} \in R_w^{(m)}\right). \tag{8}$$

This implies that the wife is differently affected by her own affective behavior, depending on the regime she is in. For instance, she may be more likely to be stuck in her affective behavior when in the extreme regimes (i.e., have an inertia parameter closer to 1) than when in the neutral regime.

Fourth, in the presentation above, the residual variance is identical across the regimes. We can also allow this variance to vary across the regimes, which implies that in certain regimes the person is allowed to be more sensitive to random environmental effects than in other regimes.

4. Parameter Estimation

We discuss two procedures for estimating the parameters of the GM model: (i) the existing procedure proposed by Gottman and Murray; and (ii) a new estimation procedure based on adapting the estimation procedure for the TAR model. In what follows, we assume that the number of regimes M_w, and the orders within each regime $P_w^{(m)}$ and $Q_w^{(m)}$ are known (for estimating the number of regimes in TAR models, see Strikholm & Teräsvirta, 2006; for estimating the orders within regimes, see De Gooijer, 2001).

4.1. Estimation Procedure Proposed by Gottman and Murray

The estimation procedure proposed by Gottman and Murray consists of two non-iterative stages. During the first stage, the parameters α_{0w} and α_{1w} are estimated using the data points $\{W_t, H_{t-1} = 0\}$. Because by definition the influence function $I_W(H_{t-1} = 0) = 0$, (2) reduces to $W_t = \alpha_w + \alpha_{1w} W_{t-1}$ for these observations. Hence, estimating α_{0w} and α_{1w} becomes a simple least squares problem (Gottman et al., 2002; Murray, 2002).

During the second stage, the parameters of the influence function are estimated, while the parameter estimates $\hat{\alpha}_{0w}$ and $\hat{\alpha}_{1w}$ obtained in the first stage are held fixed. To this end, the influence component is estimated by subtracting the estimated uninfluenced component from the observation, that is,

$$\hat{I}_W(H_{t-1}) = W_t - (\hat{\alpha}_{0w} + \hat{\alpha}_{1w} W_{t-1}). \tag{9}$$

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

Next, $\hat{I}_W(H_{t-1})$ is plotted against H_{t-1} to determine which influence function (A or B) should be used. If influence function A is considered, all the estimated influence components and their associated observed values of the influencing partner $\{\hat{I}_W(H_{t-1}), H_{t-1}\}$ are divided into cases with $H_{t-1} < 0$ and cases with $H_{t-1} > 0$. Least squares estimates of γ_w^- and γ_w^+ are obtained through regressing $\hat{I}_W(H_t)$ on H_t in each regime separately. Obtaining estimates for influence function B is slightly more complicated as it involves the estimation of two unknown thresholds. Given a particular value of τ_w^-, the estimated influences $\hat{I}_W(H_{t-1} < 0)$ are divided into those with $H_{t-1} \leq \tau_w^-$ (belonging to the negative regime) and those with $H_{t-1} > \tau_w^-$ (belonging to the neutral regime). By estimating the means of both parts of this data, the residual sum of squares can be estimated, which in turn can be used to find the optimal value of τ_w^-. Similarly, the positive threshold τ_w^+ and the constant of the positive regime δ_w^+ are estimated based on finding the optimal partitioning of $\hat{I}_W(H_{t-1} > 0)$.[2]

4.2. A New Estimation Procedure

Since the GM model is a special case of the more general TAR model, we consider the estimation procedure for the TAR model (e.g., Tong & Lim, 1980, pp. 257–258). However, in order to constrain the constant α_{0w} and the inertia α_{1w} across all regimes, we employ a slightly modified version of the standard TAR estimation procedure.

The model with influence function A contains only one threshold which is fixed at zero. Based on H_{t-1} we make two new variables: $\tilde{H}_{1,t-1}$ which has score H_{t-1} if $H_{t-1} < 0$, and score 0 otherwise; and $\tilde{H}_{2,t-1}$ which has score H_{t-1} if $H_{t-1} > 0$, and score 0 otherwise. Note that when $H_{t-1} = 0$ we have $\tilde{H}_{1,t-1} = \tilde{H}_{2,t-1} = 0$. Using these two new variables we can write

$$\begin{bmatrix} W_2 \\ W_3 \\ W_4 \\ \cdots \\ W_T \end{bmatrix} = \begin{bmatrix} 1 & W_1 & \tilde{H}_{11} & \tilde{H}_{21} \\ 1 & W_2 & \tilde{H}_{12} & \tilde{H}_{22} \\ 1 & W_3 & \tilde{H}_{13} & \tilde{H}_{23} \\ \cdots & & & \\ 1 & W_{T-1} & \tilde{H}_{1,T-1} & \tilde{H}_{2,T-1} \end{bmatrix} \begin{bmatrix} \alpha_{0w} \\ \alpha_{1w} \\ \gamma_w^- \\ \gamma_w^+ \end{bmatrix} + \begin{bmatrix} \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \cdots \\ \epsilon_T \end{bmatrix}, \tag{10}$$

and using this expression we can easily obtain the least squares estimates of all parameters (i.e., α_{0w}, α_{1w}, γ_w^-, and γ_w^+) simultaneously.

In the model with influence function B, the unknown parameters include the thresholds τ_w^- and τ_w^+. Given a particular set of threshold values, we first make two dummy variables $D_{1,t}$ and $D_{2,t}$ to identify in which of the three regimes W_t falls: The scores $D_{1,t} = 1$ and $D_{2,t} = 0$ imply that the system is in the negative regime (i.e., $H_{t-1} \leq \tau_w^-$); the scores $D_{1,t} = 0$ and $D_{2,t} = 1$ imply that the system is in the positive regime (i.e., $H_{t-1} > \tau_w^+$); the scores $D_{1,t} = 0$ and $D_{2,t} = 0$ imply that the system is in the neutral regime (i.e., $\tau_w^- < H_{t-1} \leq \tau_w^-$). We can write the following regression equation

$$\begin{bmatrix} W_2 \\ W_3 \\ W_4 \\ \cdots \\ W_T \end{bmatrix} = \begin{bmatrix} 1 & W_1 & D_{12} & D_{22} \\ 1 & W_2 & D_{13} & D_{23} \\ 1 & W_3 & D_{14} & D_{24} \\ \cdots & & & \\ 1 & W_{T-1} & D_{1T} & D_{2T} \end{bmatrix} \begin{bmatrix} \alpha_{0w} \\ \alpha_{1w} \\ \delta_w^- \\ \delta_w^+ \end{bmatrix} + \begin{bmatrix} \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \cdots \\ \epsilon_T \end{bmatrix}. \tag{11}$$

[2]Note, however, that in this procedure two extra parameters are estimated. This is necessary because, when the constant of the neutral regime is estimated, the thresholds are not known, and therefore it is not clear which observations belong to the neutral regime. Hence, the constant of the neutral regime is estimated using only part of the observations belonging to this regime, and the two extra parameters are constants for the influence components that belong to the neutral regime. Obviously, these extra parameters should be close to zero.

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

Using this expression, we can obtain the least squares estimates for the parameters α_{0w}, α_{1w}, δ_w^-, and δ_w^+, given the values for τ_w^- and τ_w^+. By varying the threshold values (and thus the values on the dummy variables $D_{1,t}$ and $D_{2,t}$), we can determine the least squares estimates of the threshold values as well as the least squares estimates of all the other model parameters simultaneously.

4.3. Standard Errors

To make inferences based on the parameter estimates, either standard errors or confidence intervals are essential. The estimation procedures described above consist of least squares estimation, and hence we can obtain the least squares standard errors for the parameter estimates. It is to be expected that the standard errors for the constant α_{w0} and the inertia parameter α_{w1} obtained with the existing estimation procedure will be larger than those obtained with the new estimation procedure because the latter uses all data to estimate these parameters, while the existing method only uses part of the data (i.e., those cases for which $H_{t-1} = 0$). For the other parameters, we do not expect differences in standard errors across the two methods.

It is important to note that neither estimation procedure results in a standard error (or confidence interval) for the estimated threshold parameters of influence function B (note that in influence function A the threshold is fixed at zero). The reason for this is that the model is estimated using linear regression, *conditional on the threshold values*. Then, by determining the sum of least squares of each possible combination of threshold values, we can determine the threshold values that minimize the sum of least squares.

The sampling distribution of the threshold parameters is nonstandard and depends on nuisance parameters such that simple inferential procedures are not available, unless one makes further assumptions (Hansen, 1997; Kapetanios, 2003). There are two alternatives for obtaining a confidence interval for threshold parameters. First, one can make use of (nonparametric) bootstrapping to construe the sampling distribution of the threshold parameters empirically. However, the dependency of the data require the use of block resampling or block subsampling techniques (Politis, 2003). Gonzalo and Wolf (2005) used a block subsampling method based on Politis and Romano (1994), and concluded that the number of occasions should be at least 500 to obtain valid confidence intervals for the threshold parameter in a TAR model. They did not investigate how large the sample size should be in case of two thresholds (i.e., for a three regime models), although it is likely to be larger than that needed for one threshold. A second alternative consists of Bayesian estimation (Chen, 1998). This approach is less likely to require large sample sizes than the block sampling techniques. The resulting empirical approximation of the posterior distribution of the threshold parameter can be used to determine a credibility interval for this parameter. We return to this issue in the discussion.

5. Simulation Studies

To compare the two estimation procedures discussed above, we perform a number of simulation studies using R. We specifically focus on: (a) the mean point estimates $\bar{\hat{x}}$ as an indication of whether the estimators are biased; (b) the sampling standard deviations $SD_{\hat{x}}$ to determine which estimation procedure is more efficient; (c) the mean standard errors $\bar{SE}_x$, which should be a good indication of the sampling standard deviation; and (d) the root mean squared errors RMSE$= \{(\bar{\hat{x}} - x)^2 + SD_{\hat{x}}^2\}^{1/2}$, which is the root of the squared bias and the variance of an estimator. In addition, we investigate whether the standard error gives a good indication of the actual sampling standard deviation.

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

5.1. *Influence Function A*

We chose the parameter values in the GM model with influence function A based on the empirical results reported by Cook et al. (1995) for what they called a validating couple and a hostile-detached couple. Because the results were very similar, we only report the results for the validating couple here (results for the hostile-detached couple are available on request from the first author). Such a couple's influence function is characterized by a positive slope when the spouse scores above zero as well as when (s)he scores below zero (i.e., γ_w^-, γ_w^+, γ_h^-, and γ_h^+ are all larger than zero). We simulated series of 150 and 300 occasions of a validating couple, representing a 15 minute and a 30 minute conversation, respectively. For both time lengths, 500 replications were simulated and then analyzed with the two estimation procedures.

From the results presented in Table 1, it can be seen that the two estimation procedures perform about equally well with regard to the average point estimate $\bar{\hat{x}}$, but that the new method performs better considering the efficiency, as can be concluded from the smaller sampling standard deviation $SD_{\hat{x}}$. This is also reflected by the larger RMSEs for the existing estimation procedure. Although the efficiency increases as the length of the series increases, the new procedure keeps outperforming the existing estimation method. Moreover, the existing method gives average standard errors that are half the size of the actual sampling standard deviation, meaning that inferences based on these estimates and standard errors will be overly optimistic. In contrast, the average standard error obtained with the new method is a reasonable representation of the corresponding sampling standard deviation.

5.2. *Influence Function B*

We simulated a GM model with influence function B and parameters that are similar to those obtained from stable couples reported by Gottman et al. (1999). Simulation results obtained with different parameter values were again quite similar and are therefore not reported here (available on request from the first author). We simulated 500 replications of 150 occasions, and another 500 replications of 300 occasions. The results are given in Table 2.

Comparing the average point estimate $\bar{\hat{x}}$ to the parameter values used for simulation, we observe more biasedness than in the previous simulation study. Especially the positive threshold and the constant of the positive regime for the husband are biased, although the bias decreases as T increases from 150 to 300. The reason for this bias is that given these parameter values, the husband is not very likely to switch to the positive regime. Since very few observations came from this regime (especially in the shorter series), it is difficult to obtain proper estimates of the parameters associated with this regime.

Again, the sampling standard deviations $SD_{\hat{x}}$ show that in general the alternative estimation procedure is more efficient than the existing one. In addition, the average standard error is a reasonable representation of the corresponding sampling standard deviation for most parameters in the alternative method. An exception is formed by the constants of the positive regimes of both the husband and the wife, but these results improve as T increases indicating that this is again related to a lack of observations within a particular regime. In contrast, the average standard errors obtained with the existing method for the constants of the negative and the positive regimes of both the husband and the wife are serious underestimates of the actual sampling standard deviations, and these do not improve much as T increases. This implies that inferences based on these standard errors will be overly optimistic. Comparing the RMSEs from both procedures leads to the conclusion that overall the new estimation procedure performs better than the existing procedure.

5.3. *Conclusion*

In sum, it can be concluded that both estimation procedures perform equally well in terms of unbiasedness, but that the new estimation procedure leads to more reliable estimates, as can

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

TABLE 1.
Simulation results for influence function A.

par.	true	$T = 150$								$T = 300$							
		$\bar{\hat{x}}$		$SD_{\hat{x}}$		$\bar{SE}_x$		RMSE		$\bar{\hat{x}}$		$SD_{\hat{x}}$		$\bar{SE}_x$		RMSE	
		exist	new	exist	new	exist	new	exist	new	exist	new	exist	new	exist	new	exist	new
α_{0w}	0.43	0.40	0.46	1.15	0.55	1.13	0.54	1.15	0.56	0.48	0.43	0.80	0.40	0.79	0.38	0.80	0.40
α_{1w}	0.14	0.13	0.12	0.33	0.08	0.30	0.08	0.33	0.08	0.13	0.13	0.21	0.06	0.20	0.06	0.21	0.06
γ_w^-	0.28	0.28	0.28	0.24	0.14	0.12	0.15	0.24	0.14	0.28	0.27	0.16	0.10	0.08	0.10	0.16	0.10
γ_w^+	0.15	0.16	0.15	0.23	0.15	0.11	0.14	0.23	0.15	0.15	0.16	0.16	0.09	0.07	0.10	0.16	0.09
α_{0h}	0.24	0.42	0.30	1.19	0.56	1.07	0.54	1.20	0.56	0.24	0.24	0.75	0.40	0.77	0.39	0.75	0.40
α_{1h}	0.37	0.33	0.36	0.30	0.07	0.26	0.07	0.30	0.08	0.37	0.36	0.18	0.05	0.18	0.05	0.18	0.05
γ_h^-	0.31	0.35	0.32	0.28	0.16	0.13	0.16	0.28	0.16	0.30	0.30	0.16	0.11	0.09	0.11	0.16	0.11
γ_h^+	0.21	0.18	0.20	0.24	0.15	0.11	0.15	0.24	0.15	0.21	0.21	0.16	0.10	0.08	0.10	0.16	0.10

Note: Simulation results for GM model with influence function A. The column heading "exist" refers to results obtained with the estimation procedure proposed by Gottman and Murray; "new" refers to the alternative estimation procedure proposed in this paper; $\bar{\hat{x}}$ is the mean of the parameter estimates obtained from 500 replications; $SD_{\hat{x}}$ is the sampling standard deviation; $\bar{SE}_x$ is the mean standard error across the 500 replications; and RMSE is the square root of the sum of the bias squared and the sampling variance.

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

TABLE 2.
Simulations results for influence function B.

par.		$T = 150$								$T = 300$							
		$\hat{x}$		$SD_{\hat{x}}$		$\bar{SE}_x$		RMSE		$\hat{x}$		$SD_{\hat{x}}$		$\bar{SE}_x$		RMSE	
		exist	new	exist	new	exist	new	exist	new	exist	new	exist	new	exist	new	exist	new
α_{0w}	0.31	0.22	0.24	1.86	0.50	1.53	0.44	1.86	0.50	0.33	0.28	1.10	0.30	1.04	0.31	1.10	0.30
α_{1w}	0.39	0.38	0.37	0.41	0.06	0.32	0.06	0.42	0.07	0.38	0.38	0.22	0.04	0.21	0.04	0.22	0.04
τ_w^-	−6.00	−6.12	−6.00	1.03	0.33	–	–	1.03	0.33	−5.99	−6.00	0.13	0.12	–	–	0.13	0.12
τ_w^+	7.00	6.27	6.40	2.00	1.70	–	–	2.13	1.80	6.90	6.85	1.11	1.05	–	–	1.12	1.06
δ_w^-	−6.20	−6.07	−6.20	3.87	0.81	0.66	0.66	3.88	0.81	−6.22	−6.17	1.84	0.56	0.41	0.56	1.84	0.56
δ_w^+	5.30	5.06	5.11	3.96	2.60	1.77	1.75	3.97	2.60	5.28	5.28	2.04	1.57	1.26	1.24	2.04	1.57
α_{0h}	0.43	0.55	0.31	1.77	0.55	1.55	0.46	1.78	0.56	0.41	0.39	1.04	0.35	1.03	0.32	1.04	0.35
α_{1h}	0.40	0.38	0.39	0.38	0.06	0.32	0.06	0.38	0.06	0.39	0.39	0.21	0.04	0.19	0.04	0.21	0.04
τ_h^-	−5.00	−5.15	−5.01	1.26	0.24	–	–	1.27	0.24	−5.00	−5.00	0.12	0.08	–	–	0.12	0.08
τ_h^+	8.00	6.66	6.80	2.51	2.40	–	–	2.84	2.69	7.52	7.69	1.97	1.77	–	–	2.03	1.80
δ_h^-	−6.50	−6.57	−6.39	3.30	0.86	0.64	0.80	3.30	0.87	−6.50	−6.47	1.76	0.56	0.40	0.56	1.76	0.56
δ_h^+	4.70	4.04	4.38	3.92	2.99	1.86	1.88	3.98	3.01	4.61	4.73	2.49	2.08	1.36	1.40	2.50	2.08

Note: Results from a GM model with influence function B.

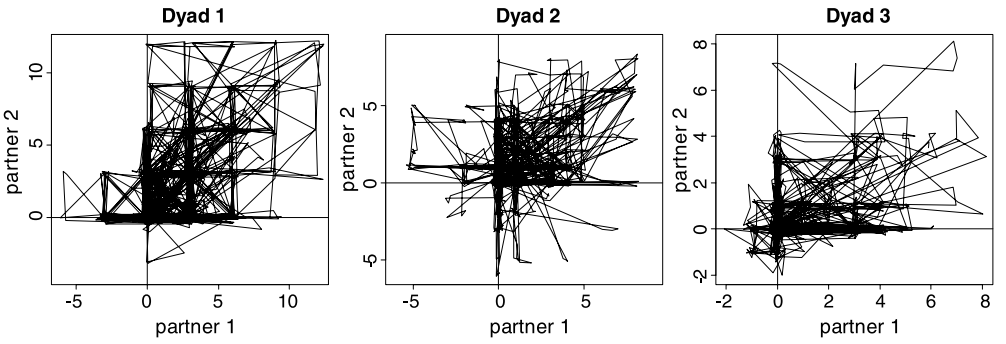


FIGURE 3.
Empirical data represented in state-space for three dyads. At each time point both partners have a score, which can be indicated as a point in state space. By connecting consecutive points over time, the current plots were created. Because the observations were made on a discrete scale, we added a small amount of noise to these data so that it becomes clearer how many observations are made at each point.

be concluded from the smaller sampling standard deviations. In addition, in comparison to the existing estimation method, the new estimation procedure results in standard errors that more closer resemble the actual sampling standard deviations. This implies that inferences based on the new estimation procedure will be more accurate than when using the existing method.

6. Empirical Illustration

Three couples were recruited at the University of Amsterdam: two heterosexual intimate couples, and a couple of female friends. Each couple engaged in three conversations: a neutral conversation (e.g., about what they have been doing in the past few days), a negative conversation (e.g., about being critical of each other), and a positive conversation (e.g., planning a trip together). The purpose of having three different conversations is that the data are likely to cover more of the state-space than if only one kind of conversation is held. This increases the possibility of obtaining enough observations from all existing regimes such that their parameters can be accurately estimated. Each conversation lasted for 15 minutes, and between subsequent conversation there was a 3 minute "cool-off" period after which the experimenter introduced the following conversation topics. Couples were allowed to choose their own topic from a list of possibilities offered by the experimenter, or to suggest a topic themselves.

Both partners were videotaped, and verbal and nonverbal behavior was coded for each spouse separately. A modified version of Gottman's SPAFF coding system was used which in combination with 2-second intervals resulted in a scale from -12 to 12. We analyzed the data of all three conversations simultaneously because we assume that, although the conversations may take place in different parts of the state-space, together they cover a large area of the state-space for this couple.[3]

We consider a variety of models, which are presented in Table 3 for the first partner (we refer to this partner as the wife: again the spousal's models are the same with only the roles of the partners reversed). All of these models can be thought of as special cases of the more general

[3]Since we do not expect the dependency between conversation to be the same as within a conversation, we do not model the first measurement occasion of any of the three conversations. That is, we only use the first observations of each conversation as a predictor and threshold variable of the second observation, but we do not consider it as a realization of the dependent variable.

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

TABLE 3.
Model selection for three dyads.

	Model	M_w	dyad 1		dyad 2		dyad 3	
			1	2	1	2	1	2
1a:	$W_t = \sum_{m=1}^{M_w} \{\alpha_{0w} + \alpha_{1w} W_{t-1} + \gamma_w^{(m)} H_{t-1} + \sigma \epsilon_t\} I(H_{t-1} \in R_w^{(m)})$	2	4055	3929	3179	3298	2108	1414
1b:	$W_t = \sum_{m=1}^{M_w} \{\alpha_{0w} + \alpha_{1w} W_{t-1} + \gamma_w^{(m)} H_{t-1} + \sigma^{(m)} \epsilon_t\} I(H_{t-1} \in R_w^{(m)})$	2	4016	3786	3098	3209	2024	1281
2a:	$W_t = \sum_{m=1}^{M_w} \{\alpha_{0w}^{(m)} + \alpha_{1w} W_{t-1} + \sigma \epsilon_t\} I(H_{t-1} \in R_w^{(m)})$	2	4046	3941	3172	3286	2106	1409
		3	4051	3931	3171	3291	2112	1415
2b:	$W_t = \sum_{m=1}^{M_w} \{\alpha_{0w}^{(m)} + \alpha_{1w} W_{t-1} + \sigma^{(m)} \epsilon_t\} I(H_{t-1} \in R_w^{(m)})$	2	*4009*	3796	*3090*	*3198*	2023	1274
		3	4012	*3772*	3099	3198	*2007*	*1258*
3a:	$W_t = \sum_{m=1}^{M_w} \{\alpha_{0w}^{(m)} + \alpha_{1w} W_{t-1} + \gamma_w^{(m)} H_{t-1} + \sigma \epsilon_t\} I(H_{t-1} \in R_w^{(m)})$	2	4057	3936	3171	3294	2115	1421
		3	4069	3947	3180	3303	2129	1430
3b:	$W_t = \sum_{m=1}^{M_w} \{\alpha_{0w}^{(m)} + \alpha_{1w} W_{t-1} + \gamma_w^{(m)} H_{t-1} + \sigma^{(m)} \epsilon_t\} I(H_{t-1} \in R_w^{(m)})$	2	4021	3793	3104	3209	2031	1287
		3	4032	3777	3112	3214	2023	1275

Note: Bayesian information criteria for a variety of models considered for each of the two partners of the three dyads. Smaller values indicate a better model. The smallest value per person is given in italics.

expression

$$W_t = \sum_{m=1}^{M_w} \left\{ \alpha_{0w}^{(m)} + \alpha_{1w} W_{t-1} + \gamma^{(m)} H_{t-1} + \sigma^{(m)} \epsilon_t \right\} I\left(H_{t-1} \in R_w^{(m)}\right), \tag{12}$$

which in turn is a special case of the TAR model specified in (1).

The first model is related to the GM model with influence function A, and is obtained from (12) by having $M_w = 2$, $\alpha_{0w}^{(m)} = \alpha_{0w}$, and $\alpha_{1w}^{(m)} = \alpha_{1w}$. In contrast to the actual GM model with influence function A, we do not fix the threshold at zero, but estimate it as a free parameter. Two versions of this model are considered: a model with equal residual variances in both regimes, and a model with differing residual variances (see Table 3, models 1a and 1b).

The second model is based on the GM model with influence function B, and is obtained from the more general model in (12) by letting $\gamma_w^{(m)} = 0$ for all m. We consider both $M_w = 2$ and $M_w = 3$; the latter is the actual GM model with influence function B (with $\alpha_{0w}^{(2)} = \alpha_{0w}$, $\alpha_{0w}^{(1)} = \alpha_{0w} + \delta_w^-$, and $\alpha_{0w}^{(3)} = \alpha_{0w} + \delta_w^+$). Again we consider these models both with equal residual variances and unequal residual variances across the regimes (see models 2a and 2b in Table 3, respectively).

The third model can be thought of as a combination of the two influence functions proposed by Gottman and his colleagues. It consists of a step function as in influence function B, but within each regime the partner's score is included as it is in influence function A. It is obtained from the general expression in (12) by having $M_w = 2$ and $M_w = 3$. Again we look at models with equal and with unequal residual variances (see models 3a and 3b in Table 3).

Extended simulations studies have shown that the Bayesian Information Criterion (BIC; Schwarz, 1978) is an appropriate measure for determining the number of regimes of a TAR model (Gonzalo & Pitarakis, 2002; Strikholm & Teräsvirta, 2006).[4] Hence, we make use of the BIC to decide which of the models described above is the most appropriate for the data of the 3 dyads. In Table 3, the BIC for each person is given, with the smallest BIC value printed in italics. It can be seen that for all individuals the model based on influence function B with unequal residual variances is most appropriate, with either two or three regimes. Hence, we can conclude that a step function is most appropriate to describe the influence of one partner on the other, and that the partner's score does not help to further predict the observations within the regimes.

In Table 4, the parameter estimates and corresponding standard errors for the most appropriate model per person are presented. In addition, the proportion of explained variance (R^2) is given.[5] Based on the parameter estimates, we determined the nullclines for each of the couples, and depicted these in Figure 4. From this it can be seen that Dyad 1 and Dyad 3 have only one influenced stable steady state, while Dyad 2 has two influenced stable steady states (which are very close to each other). All these influenced stable steady states are positioned in the positive quadrant, meaning that all three dyads are attracted toward a positive way of interacting with each other. This is not surprising, as these were nonclinical couples (i.e., they were not seeking

[4]Using the log likelihood ratio test for nested models is not an option here because this statistic has an unknown distribution when comparing a model with M regimes to a model with $M + 1$ regimes, unless the thresholds are known in advance (e.g., Chan & Tong, 1990). The reason for this is that the Mth threshold is a nuisance parameter that is absent under the null hypothesis.

[5]In time series analysis models, there are no fit statistics such as in regression analysis or structural equation modeling to evaluate a model. The reason for this is that the quality of a time series models consists of not only explaining as much variance as possible using previous observations, but also explaining as much of the covariance between observations at different occasions. One way to think about a time series is that it is a single draw from a T-variate distribution, and one can actually model a time series as such using structural equation modeling that allow raw data likelihood optimization (see Hamaker, Dolan, & Molenaar, 2003). However, because the number of "variables" (i.e., occasions) is larger than the number of cases (i.e., $n = 1$), it is not possible to estimate the saturated model, and as a result it is not possible to compute a chi-square test statistic.

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

TABLE 4.
Results for three dyads.

dyad	partner	α_0	α_1	δ^-	δ^+	τ^-	τ^+	R^2
1	1	0.53 (0.06)	0.68 (0.02)	–	0.93 (0.22)	–	7	0.51
	2	0.21 (0.15)	0.73 (0.02)	0.05 (0.17)	0.74 (0.17)	0	2	0.61
2	1	0.19 (0.04)	0.77 (0.02)	–	0.16 (0.08)	–	1	0.60
	2	0.25 (0.04)	0.73 (0.02)	–	0.26 (0.09)	–	1	0.55
3	1	0.30 (0.07)	0.75 (0.02)	−0.15 (0.07)	0.16 (0.16)	0	3	0.61
	2	0.18 (0.05)	0.77 (0.02)	−0.12 (0.05)	0.15 (0.10)	0	3	0.64

Note: Parameter estimates obtained with a nonparametric bootstrap procedure for the constant α_0, the inertia α_1, steps δ^- and δ^+ of the negative and positive regime, respectively, thresholds τ^- and τ^+. Standard errors are given in parentheses. The proportion of explained variance is also given in the last column. For individuals that were identified by a model with only two regimes, these are referred to as the neutral and positive regime.

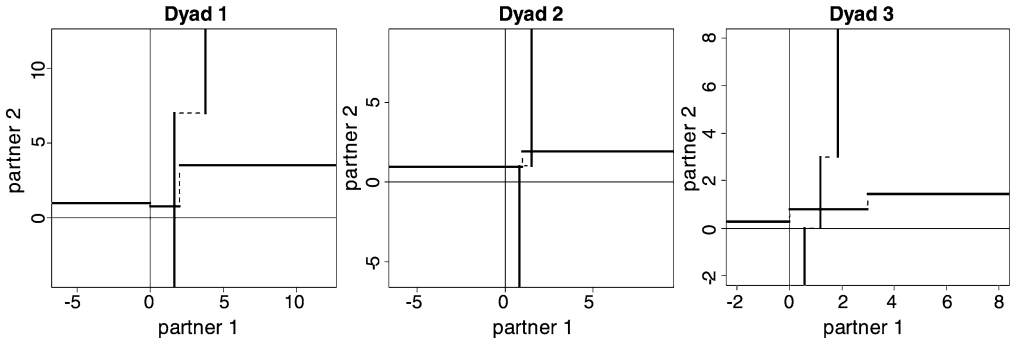


FIGURE 4.
Nullclines for the three dyads based on the models in Table 4. Vertical lines belong to partner 1 and divide the state space into a left area in which there is a tendency to move to the right (i.e., an increase in positive affective behavior of partner 1), and a right area in which there is a tendency to move to the left (i.e., a decrease in positive affective behavior of partner 1). Horizontal lines belong to partner 2 and divide the state-space into an area above the nullcline in which there is a tendency to move down (i.e., a decrease in positive affective behavior of partner 2), and an area below the nullcline in which there is a tendency to move upwards (i.e., an increase in positive affective behavior of partner 2). Intersections of the nullclines indicate attractors, that is, states to which the system is attracted.

help). Moreover, the uninfluenced steady states are all positive, meaning that these individuals are likely to behave positively when the partner does not influence them. Finally, all inertia parameters lie between 0.68 and 0.77, meaning that all individuals are characterized by a fairly strong tendency to remain in a particular state.

7. Discussion

The aim of this paper was to tackle some unresolved issues regarding parameter estimation and model selection in the context of the GM model. Through showing that the GM model is a special case of the TAR model, a new estimation procedure was proposed, which proved to be superior to the existing method in that it leads to more efficient estimates, and to standard errors that represent the actual sampling standard deviation quite well. In addition, recognizing

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

that the GM model is a TAR model allowed us to make use of existing knowledge from the TAR literature regarding model selection, which we illustrated in the empirical example. Finally, placing the GM model in the class of TAR models makes it easy to formulate alternative models, such as the ones we considered in our illustration.

While some of the most spectacular results obtained with the GM model come from making cross-dyadic comparisons (e.g., Cook et al., 1995; Gottman et al., 1999), we have not focused on this issue in the current paper. Obviously, before cross-dyadic comparisons can be made, the researcher has to decide on the actual model at the within-dyad level, and consequently we chose to focus on solutions for problems associated with this cardinal first step. However, a few comments about cross-dyadic comparisons are in place here.

To compare dyads and to determine whether certain characteristics of the interaction patterns are predictive of, for instance, future marital status and satisfaction, Gottman and his colleagues have embraced the following procedure. First, they estimate the parameters of the GM model for each dyad separately using their two-stage procedure discussed earlier in this paper, and subsequently they use these parameter estimates and functions thereof (i.e., the uninfluenced and influenced steady states), as variables in a cross-dyadic comparison. These cross-dyadic comparisons may consist of computing correlations between the dyadic features and various outcome variables (such as marital satisfaction and health, see Cook et al., 1995), or of performing ANOVAs on the parameter estimates using group membership (e.g., together, separated or divorced) as the independent variable (e.g., Gottman et al., 1999).

This approach, however, ignores the fact that there are differences in the precision with which the dyadic parameters have been estimated. One way to improve this approach is to make use of shrinkage estimators (also known as empirical Bayes estimators). Shrinkage estimators are often used in meta-analyses, in which parameters from different studies are combined, while differences in precision are also taken into account (cf. Normand, 1999; Stijnen, 2000). However, to obtain shrinkage estimators, standard errors are needed for all parameters. This poses a problem for TAR models, as the current estimation procedures do not generate standard errors for the threshold parameters. Moreover, it is difficult to determine the standard errors of quantities that are functions of model parameters, such as the influenced and uninfluenced steady states.

A more sophisticated approach to combining the information of multiple dyads consists of developing a multilevel model based on the GM model and its extensions. Clearly, cross-dyadic comparisons fit within the framework of multilevel modeling, in which repeated measures are nested within dyads. Since TAR models are closely related to change-point models (cf. Tsay, 1998, p. 1189), multilevel extensions of the latter could prove useful for this purpose (see Wang & McArdle, 2008).

Appendix

The GM model with random shocks can be expressed as

$$W_t = \alpha_{0w} + \alpha_{1w} W_{t-1} + I_W(H_{t-1}) + e_{wt}, \tag{13}$$

where e_{wt} is normally distributed with mean zero and variance σ_w^2. The wife's nullcline is the function where her behavior is in equilibrium, meaning that her expected behavior does not change. To find this function, we first write

$$\begin{aligned} \mathrm{E}[W_t] &= \mathrm{E}\big[\alpha_{0w} + \alpha_{1w} W_{t-1} + I_W(H_{t-1}) + e_{wt}\big] \\ &= \alpha_{0w} + \alpha_{1w}\mathrm{E}[W_{t-1}] + I_W(H_{t-1}). \end{aligned}$$

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

Making use of the fact that (if the wife's behavior is in equilibrium), $\mathrm{E}[W_t] = \mathrm{E}[W_{t-1}]$, we can write

$$\begin{aligned}\alpha_{0w} + I_W(H_{t-1}) &= \mathrm{E}[W_t] - \alpha_{1w}\mathrm{E}[W_{t-1}] \\ &= (1-\alpha_{1w})\mathrm{E}[W_t],\end{aligned}$$

and we thus obtain

$$\mathrm{E}[W_t] = \frac{\alpha_{0w} + I_W(H_{t-1})}{1-\alpha_{1w}}.$$

The latter expression shows that the nullcline is a linear transformation of the influence function, so that it will have the same shape as the influence function. Similarly, the nullcline for the husband, for which $\mathrm{E}[H_t] = \mathrm{E}[H_{t-1}]$, can be expressed as

$$\mathrm{E}[H_t] = \frac{\alpha_{0h} + I_H(W_{t-1})}{1-\alpha_{1h}}.$$

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

Bisconti, T., Bergeman, C. S., & Boker, S. M. (2004). Emotional well-being in recently bereaved widows: A dynamical system approach. *Journal of Gerontology, Series B: Psychological Sciences and Social Sciences*, *59*, 158–167.

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York: Cambridge University Press.

Chan, K. S., & Tong, H. (1990). On likelihood ratio tests for threshold autoregression. *Journal of the Royal Statistical Society*, *52*, 469–476.

Chen, C. W. S. (1998). A Bayesian analysis of generalized threshold autoregressive models. *Statistical and Probability Letters*, *40*, 15–22.

Chow, S.-M., Ferrer, E., & Nesselroade, J. R. (2007). An unscented Kalman filter approach to the estimation of nonlinear dynamic system models. *Multivariate Behavioral Research*, *42*, 283–321.

Clayton, K. (1997). *Basic concepts in nonlinear dynamics and chaos*. Workshop presented at the society for chaos theory in psychology and the life sciences meeting.

Cook, J., Tyson, R., White, R. R., Gottman, J. M., & Murray, J. (1995). Mathematics of marital conflict: Qualitative dynamic mathematical modeling of marital interaction. *Journal of Family Psychology*, *9*, 110–130.

De Gooijer, J. G. (2001). Cross-validation criteria for SETAR model selection. *Journal of Time Series Analysis*, *22*, 267–281.

Fan, J., & Yao, Q. (2003). *Nonlinear time series: Nonparametric and parametric methods*. New York: Springer.

Gonzalo, J., & Pitarakis, J.-Y. (2002). Estimation and model selection based inference in single and multiple threshold models. *Journal of Econometrics*, *110*, 319–352.

Gonzalo, J., & Wolf, M. (2005). Subsampling inference in threshold autoregressive models. *Journal of Econometrics*, *127*, 201–224.

Gottman, J. M., Coan, J., Carrere, S., & Swanson, C. (1998). Predicting marital happiness and stability from newlywed interactions. *Journal of Marriage and the Family*, *60*, 5–22.

Gottman, J. M., Levenson, R. W., Swanson, C., Swanson, K., Tyson, R., & Yoshimoto, D. (2003). Observing gay, lesbian, and heterosexual couple's relationships: Mathematical modeling of conflict interaction. *Journal of Homosexuality*, *45*, 65–91.

Gottman, J. M., McCoy, K., Coan, J., & Collier, H. (1996). The specific affect coding system SPAFF. In J. M. Gottman (Ed.), *What predicts divorce? The measures* (pp. 1–169). Hillsdale: Lawrence Erlbaum Associates.

Gottman, J. M., Murray, J. D., Swanson, C. C., Tyson, R., & Swanson, K. R. (2002). *The mathematics of marriage: Dynamic nonlinear models*. Cambridge: MIT Press.

Gottman, J. M., Swanson, C. C., & Murray, J. D. (1999). The mathematics of marital conflict: Dynamic mathematical nonlinear modeling of newlywed marital interaction. *Journal of Family Psychology*, *13*, 3–19.

Granger, C. W. J., & Andersen, A. P. (1978). *An introduction to bilinear time series models*. Göttingen: Vandenhoeck und Ruprecht.

Granic, I., & Hollenstein, T. (2003). Dynamic system methods for models of developmental psychopathology. *Development and Psychopathology*, *15*, 641–669.

Guastello, S. J. (1997). Science evolves: An introduction to nonlinear dynamics, psychology, and life sciences. *Nonlinear Dynamics, Psychology, and Life Sciences*, *1*, 1–6.

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2003). ARMA-based SEM when the number of time points T exceeds the number of cases N: Raw data maximum likelihood. *Structural Equation Modeling*, *10*, 352–379.
Hansen, B. E. (1997). Inference in TAR models. *Studies in Nonlinear Dynamics and Econometrics*, *2*, 1–14.
Kapetanios, G. (2003). *Using extraneous information and GMM to estimate threshold parameters in TAR models* (U of London Queen Mary Economics Working Paper No. 494). Available at SSRN: http://ssrn.com/abstract=425380 or doi:10.2139/ssrn.425380.
Murray, J. D. (2002). *Mathematical biology: I. An introduction* (3 ed.). New York: Springer.
Normand, S.-L. (1999). Tutorial in biostatistics. Meta-analysis: Formulating, evaluating, combining and reporting. *Statistics in Medicine*, *18*, 321–359.
Olthof, T., Kunnen, E. S., & Boom, J. (2000). Simulating mother-child interaction: Exploring two varieties of a non-linear dynamic system approach. *Infant and Child Development*, *9*, 33–60.
Politis, D. N. (2003). The impact of bootstrap methods on time series analysis. *Statistical Science*, *18*, 219–230.
Politis, D. N., & Romano, J. P. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *The Annals of Statistics*, *22*, 2031–2050.
Schiepek, G. (2003). A dynamic system approach to clinical case formulation. *European Journal of Psychological Assessment*, *19*, 175–184.
Schwartz Gottman, J. (2004). *The marriage clinic casebook*. New York: W.W. Norton and Co.
Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464.
Shoda, Y., Tiernan, S. L., & Mischel, W. (2002). Personality as a dynamic system: Emergence of stability and distinctiveness from intra- and interpersonal interactions. *Personality and Social Psychology Review*, *6*, 316–325.
Stijnen, T. (2000). Tutorial in biostatistics. meta-analysis: Formulating, evaluating, combining and reporting. *Statistics in Medicine*, *19*, 753–761.
Strikholm, B., & Teräsvirta, T. (2006). A sequential procedure for determining the number of regimes in a threshold autoregressive model. *Econometrics Journal*, *9*, 472–491.
Thelen, E., & Smith, L. B. (1994). *A dynamic system approach to the development of cognition and action*. Cambridge: MIT Press.
Tong, H., & Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society, B*, *42*, 245–292.
Tsay, R. S. (1998). Testing and modeling multivariate threshold models. *Journal of the American Statistical Association*, *93*, 1188–1202.
Vallacher, R. R., & Nowak, A. (1997). The emergence of dynamical social psychology. *Psychological Inquiry*, *8*, 73–99.
Vallacher, R. R., Read, S. J., & Nowak, A. (2002). The dynamical perspective in personality and social psychology. *Personality and Social Psychology*, *4*, 264–273.
Van der Maas, H. L. J., & Molenaar, P. C. M. (1992). Stagewise cognitive development: An application of catastrophe theory. *Psychological Review*, *99*, 395–417.
Van der Maas, H. L. J., & Raijmakers, M. E. J. (2000). A phase transition model for mother-child interaction: Comment on Olthof et al., 2000. *Infant and Child Development*, *9*, 75–83.
Van Geert, P., & Van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development*, *25*, 340–374.
Wang, L., & McArdle, J. J. (2008). A simulation study comparison of Bayesian estimate with conventional methods for estimating unknown change points. *Structural Equation Modeling*, *15*, 52–74.
Warren, K. (2002). Thresholds and the abstinence violation effect: A nonlinear dynamic model of the behaviors of intellectually disabled sex offenders. *Journal of Interpersonal Violence*, *17*, 1198–1217.
Warren, K., Hawkins, R. C., & Sprott, J. C. (2003). Substance abuse as a dynamical disease: Evidence and clinical implications of nonlinearity in a time series of daily alcohol consumption. *Addictive Behaviors*, *28*, 369–374.
Witkiewitz, K., Van der Maas, J. L., Hufford, M. R., & Marlatt, G. A. (2007). Nonnormality and divergence in posttreatment alcohol use: Reexamining the Project MATCH data another way. *Journal of Abnormal Psychology*, *116*, 378–394.

Manuscript Received: 27 MAY 2008
Final Version Received: 8 DEC 2008
Published Online Date: 25 FEB 2009

https://doi.org/10.1007/s11336-009-9113-4 Published online by Cambridge University Press