Supplemental Materials:
On Subgrouping Continuous Processes in Discrete Time

Jonathan J. Park, Zachary Fisher, Sy-Miin Chow, and Peter C. M. Molenaar

Department of Human Development and Family Studies, The Pennsylvania State University

Supplemental Materials:

On Subgrouping Continuous Processes in Discrete Time

Simulation Design

In empirical- and simulation-based studies, VAR-based approaches have been used to identify meaningful subgroups with unique structural differences in their discrete-time network dynamics (e.g., $\mathbf{\Phi}$, $\mathbf{A}$, or $\mathbf{PCC}$; Bulteel et al., 2016; Henry et al., 2019; Price et al., 2017). However, the use of intensive longitudinal data is wrought with questions regarding the optimal alignment of time-scales and sampling rates (Ram & Diehl, 2014; Ryan & Hamaker, 2022). For example, these issues arise in studies of daily affect (Wright et al., 2019) where emotions unfold and influence each other faster than data are typically collected and fMRI where neuronal activity occurs significantly faster than the imaging rate (James et al., 2019).

To address these questions, we simulated 2 subgroups of individuals of sample size $N_{total} = 50$ $(n_{g1} = 25; n_{g2} = 25)$ using a 4-variate Ornstein-Uhlenbeck (OU) model (see Oravecz & Tuerlinckx, 2011). The OU model is considered the continuous-time analogue for the discrete-time VAR(1) and is presented in Equation 1 in the main text as a special case of Equation **??**.

$$\mathrm{d}\boldsymbol{\eta}(t) = \mathbf{B}[\boldsymbol{\mu} - \boldsymbol{\eta}(t)]\,\mathrm{d}t + \mathbf{G}\mathrm{d}\mathbf{W}(t) \tag{1}$$

here, $\mathbf{A}$ in Equation **??** is represented as $-\mathbf{B}$ and $\mathbf{b} = \mathbf{B}\boldsymbol{\mu}$, where $\mathbf{B}$ is the drift matrix and $\boldsymbol{\mu}$ is a vector of intercepts. Both subgroups shared four auto-regressive and one cross-process effect ($B_{1,1}, B_{2,2}, B_{3,3}, B_{4,4}, B_{3,1}$) and differed on two cross-process effects, one in terms of sign only ($B_{1,2}$), and one in terms of path location ($B_{2,4}$ in subgroup 1 and $B_{4,3}$ in subgroup 2).

We examined the influence of effect size, specifically, separation between the two subgroups in drift matrix coefficients under a small and large effect size conditions,

with drift matrices: $\mathbf{B}_{g1}$ and $\mathbf{B}_{g2}$.

$$\mathbf{B_{g1}} = \begin{bmatrix} 0.50 & \sigma & 0.00 & 0.00 \\ 0.00 & 0.50 & 0.00 & -\sigma \\ -\alpha & 0.00 & 0.50 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.50 \end{bmatrix}, \mathbf{B_{g2}} = \begin{bmatrix} 0.50 & -\sigma & 0.00 & 0.00 \\ 0.00 & 0.50 & 0.00 & 0.00 \\ -\alpha & 0.00 & 0.50 & 0.00 \\ 0.00 & 0.00 & -\sigma & 0.50 \end{bmatrix}. \quad (2)$$

where σ–subgroup specific paths–and α–common paths–took on values of 0.30 or 0.60 in the small and large effect size conditions, respectively.

In addition, we repeated the evaluation of effect size using an alternative set of drift coefficients, namely, with cross-process coefficients taking on values of 0.90 and 1.20. This yielded two additional sets of coefficients that mirrored the first evaluation in terms of the separation between subgroups (i.e., "effect size"), but both subgroups' coefficients deviated further from zero and were more prone to show sustained deviations from, or slower return to the system's intercepts. In the other words, this second set of drift coefficients generated dynamics that were closer to the boundary of being unstable (Lütkepohl, 2005).

We considered time-series lengths of $T = 14$ and $T = 100$ to emulate a 2-week long experiment, as well as a sample-size more frequently seen in the applied time-series literature (De Vos et al., 2017). Once generated, the continuous data from the OU models were subsampled at intervals of $0.1s, 0.5s, 1.0s$, and $10.0s$. To obtain the target T, we sampled the last 14 or 100 time points of the generated series to ensure that the effects of any initial (transient) dynamics on the data were discarded.

References

Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2016). Clustering vector autoregressive models: Capturing qualitative differences in within-person dynamics. *Frontiers in Psychology*, *7*, 1540.

De Vos, S., Wardenaar, K. J., Bos, E. H., Wit, E. C., Bouwmans, M. E., & De Jonge, P. (2017). An investigation of emotion dynamics in major depressive disorder patients and healthy persons using sparse longitudinal networks. *PLoS One*, *12*(6), e0178586.

Henry, T. R., Feczko, E., Cordova, M., Earl, E., Williams, S., Nigg, J. T., Fair, D. A., & Gates, K. M. (2019). Comparing directed functional connectivity between groups with confirmatory subgrouping gimme. *NeuroImage*, *188*, 642–653.

James, O., Park, H., & Kim, S.-G. (2019). Impact of sampling rate on statistical significance for single subject fmri connectivity analysis. *Human brain mapping*, *40*(11), 3321–3337.

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

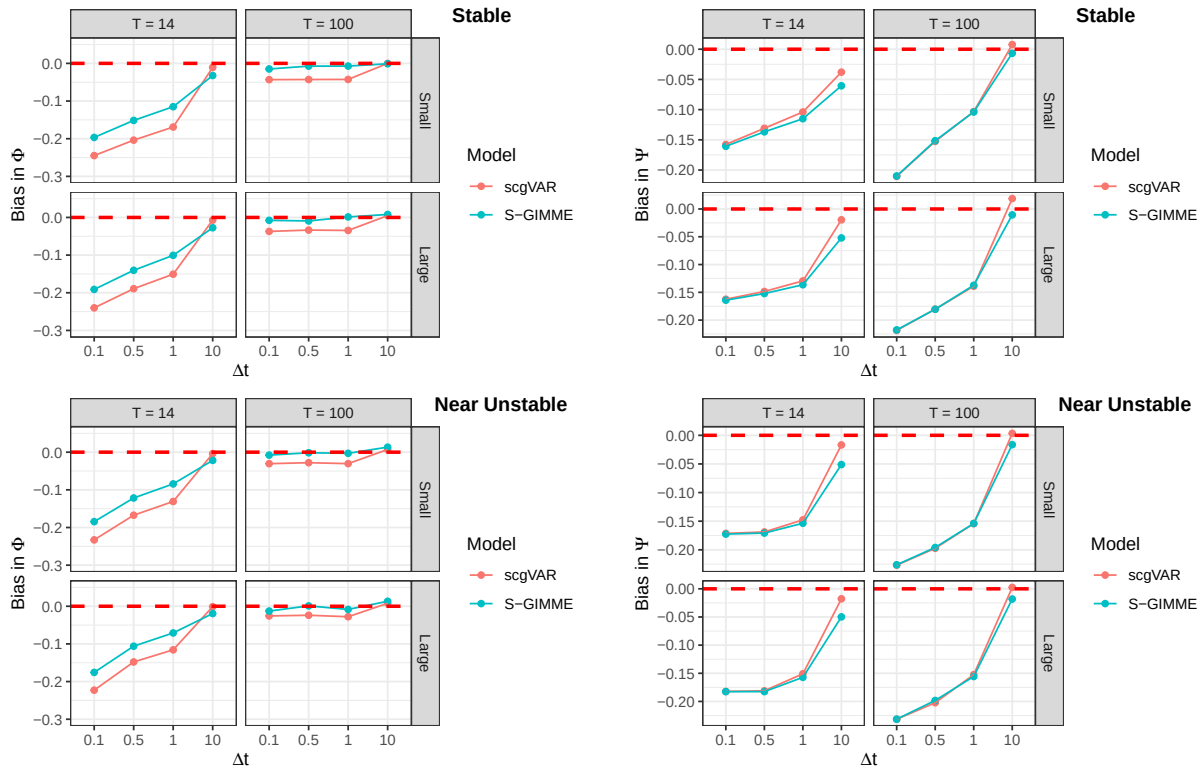
Oravecz, Z., & Tuerlinckx, F. (2011). The linear mixed model and the hierarchical ornstein–uhlenbeck model: Some equivalences and differences. *British Journal of Mathematical and Statistical Psychology*, *64*(1), 134–160.

Price, R. B., Lane, S., Gates, K., Kraynak, T. E., Horner, M. S., Thase, M. E., & Siegle, G. J. (2017). Parsing heterogeneity in the brain connectivity of depressed and healthy adults during positive mood. *Biological psychiatry*, *81*(4), 347–357.

Ram, N., & Diehl, M. (2014). Multiple-time-scale design and analysis: Pushing toward real-time modeling of complex developmental processes. *Handbook of intraindividual variability across the life span* (pp. 328–343). Routledge.

Ryan, O., & Hamaker, E. L. (2022). Time to intervene: A continuous-time approach to network analysis and centrality. *Psychometrika*, *87*(1), 214–252.

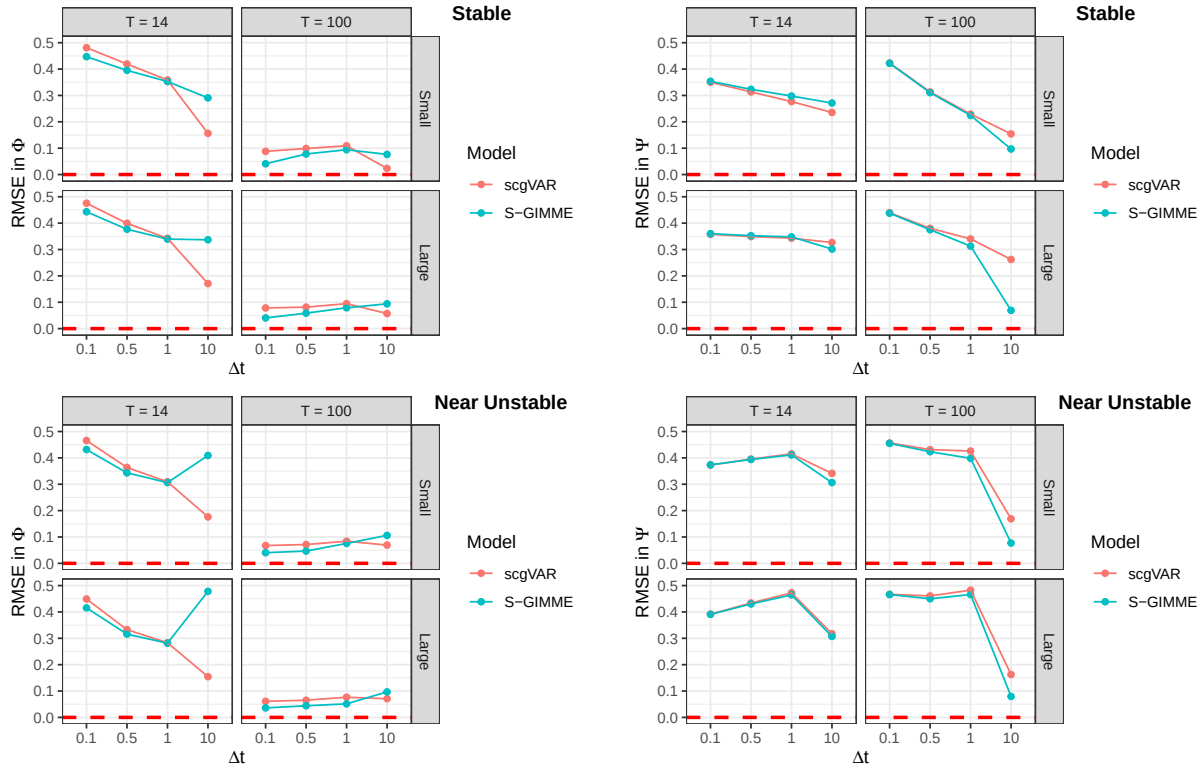
Wright, A. G., Gates, K. M., Arizmendi, C., Lane, S. T., Woods, W. C., & Edershile, E. A. (2019). Focusing personality assessment on the person: Modeling general, shared, and person specific processes in personality and psychopathology. *Psychological Assessment*, *31*(4), 502.



(a) Biases for Φ-parameters across Δt

(b) Biases for $\mathbf{\Psi}$-parameters across Δt

Figure 1. Biases of non-zero parameters for both S-GIMME and the scgVAR when transformed to the VAR metric. Biases were calculated as: $\text{Bias}(\theta) = \frac{1}{H}\sum_{h=1}^{H}\left(\hat{\theta}_h - \theta\right)$. θ and $\hat{\theta}$ represent the parameters and their estimates, H is the number of simulations.



(a) RMSE for Φ-parameters across Δt

(b) RMSE for $\boldsymbol{\Psi}$-parameters across Δt

Figure 2. RMSE's of non-zero parameters for both S-GIMME and the scgVAR when transformed onto the VAR metric. RMSEs were calculated as: $RMSE(\theta) = \sqrt{\frac{1}{H}\sum_{h=1}^{H}(\hat{\theta}_h - \theta)^2}$