Published in final edited form as:

Psychometrika. 2017 December ; 82(4): 875–903. doi:10.1007/s11336-017-9587-4.

Bayesian Sensitivity Analysis of Dynamic Factor Analysis Models with Nonparametric Prior and Possible Nonignorable Missingness

Niansheng Tang^{1,†}, Sy-Miin Chow^{2,†}, Joseph G. Ibrahim³, and Hongtu Zhu³

¹Department of Statistics, Yunnan University, Kunming, P. R. China

²Department of Human Development and Family Studies, Pennsylvania State University, U.S.A

³Department of Biostatistics, University of North Carolina at Chapel Hill, U.S.A

Abstract

Many psychological concepts are unobserved and usually represented as latent factors apprehended through multiple observed indicators. When multiple-subject multivariate time series data are available, dynamic factor analysis models with random effects offer one way of modeling patterns of within- and between-person variations by combining factor analysis and time series analysis at the factor level. Using the Dirichlet Process (DP) as a nonparametric prior for individual-specific time series parameters further allows the distributional forms of these parameters to deviate from commonly imposed (e.g., normal or other symmetric) functional forms, arising as a result of these parameters' restricted ranges. Given the complexity of such models, a thorough sensitivity analysis is critical but computationally prohibitive. We propose a Bayesian local influence method that allows for simultaneous sensitivity analysis of multiple modeling components within a single fitting of the model of choice. Five illustrations and an empirical example are provided to demonstrate the utility of the proposed approach in facilitating the detection of outlying cases and common sources of misspecification in dynamic factor analysis models, as well as identification of modeling components that are sensitive to changes in the DP prior specification.

Keywords

Bayesian local influence; Bayesian perturbation manifold; Dirichlet process prior; Nonignorable missing data; Nonlinear dynamic factor analysis model; sensitivity analysis

The past two decades have evidenced a gradual but stable increase in the usage of Bayesian methods in dynamic modeling of latent (unobserved) variables. The flexibility afforded by Bayesian and the related Markov chain Monte Carlo (MCMC) techniques holds tremendous promise but also warrants some caution in model inference and interpretations. One distinct characteristic of Bayesian data analysis in comparison to frequentist approaches is the need

Correspondence concerning this article can be addressed to Sy-Miin Chow, the Pennsylvania State University, 413 Biobehavioral Health Building, University Park, PA 16802 or by symiin@psu.edu.

[†]These two authors contributed equally to the work.

to consider prior distributions. From conjugate priors that yield posterior distributions of known analytic forms (for examples see Lee, 2007) to nonparametric priors that afford more flexibility while imposing some limited prior beliefs (e.g., Ansari and Iyengar, 2006; Chow et al., 2011; Duncan and MacEachern, 2008; Lee et al., 2007; Navarro et al., 2006), Bayesian models have assumed increasingly complex forms in recent years. As such, it is imperative to evaluate the sensitivity of the estimation results to changes in the components implicated in the analysis, including assumptions concerning the nonparametric prior distributions, as well as other prior and sampling distributions in the model (Berger, 1994).

There is an extensive literature on sensitivity analysis in the frequentist framework, especially in contexts where some of the modeling assumptions are difficult to test or falsify empirically, such as in the modeling of missing data problems (Copas and Eguchi, 2005; Little and Rubin, 2002; Zhu et al., 2007; van Steen et al., 2001; Troxel et al., 1998; Jansen et al., 2006; Verbeke et al., 2001; Shi et al., 2009; Hens et al., 2006; Daniels and Hogan, 2008). With the inclusion of priors and the enriched opportunities for more complex modeling options afforded by the Bayesian framework and the associated MCMC computational techniques, the need for Bayesian sensitivity analysis is well-known and well-articulated (Daniels and Hogan, 2008). Indeed, many methods have been developed to compute analytic or numerical derivatives of the posterior distributions or related influence measures with respect to changes in a limited subset of modeling components, such as the priors (Dey et al., 1996; Gustafson, 1996a,b; Berger, 1994; Millar and Stewart, 2007; van der Linde, 2007; Lavine, 1991), as measures of sensitivity. Methods involving the use of analytic derivative methods are not applicable in cases where the posterior distributions of interest do not have closed-form expressions. Even though numerical derivatives may be used as a viable alternative in some scenarios, inference and computational costs can still be prohibitive when researchers are interested in perturbing multiple (often interrelated) elements in the model. A recent local influence approach proposed by Zhu et al. (2011) offers a comprehensive framework for attaining these aims with a wide range of models. However, applications of this approach remain scarce in the psychometric literature, and its utility as a tool for detecting model misspecifications and influential outliers in dynamic factor analysis models (DFA) and other related variations remains unexplored.

Dynamic Factor Analysis (DFA) Models

Dynamic factor analysis (DFA) models with categorical items and random effects are one example of models that require thorough sensitivity analysis. By combining item factor analysis and time series analysis, such dynamic factor analysis models accomplish the goals of item consolidation and representation of within- and between-person variations within a single model (Browne and Nesselroade, 2005; Engle and Watson, 1981; Forni et al., 2005; Geweke, 1977; Geweke and Singleton, 1981; Molenaar, 1985; Nesselroade et al., 2002). Substantively, dynamic factor analysis models have been used to study psychophysiological processes (Molenaar, 1994), affective processes of couples (Ferrer and Nesselroade, 2003), emotions of patients with Parkinson's disease (Chow et al., 2004), emotion process after a romantic breakup (Sbarra, 2006) and relationship between children's perceived control and school performance (Musher-Eizenman et al., 2002).

More recently, Chow et al. (2011) proposed using the Dirichlet process (DP) prior (Antoniak, 1974; Ferguson, 1973) as a nonparametric prior for the random effects in the time series parameters in a DFA model to account for such parameters' restrictions in range and possible asymmetry in distributional forms. The functional relations among the latent factors are allowed to be nonlinear. Despite the model's theoretical appeal, multiple conditional distributions in the proposed model do not have closed-form solutions. Metropolis-Hasting algorithms were embedded within a Gibbs sampler to enable estimation of modeling parameters and other unknown components (e.g., latent factors) with ordinal data. The complexity of this model calls for a thorough sensitivity analysis – a task that is unfortunately hampered by the difficulties in simultaneously evaluating the tenability of multiple assumptions and the effects of outliers.

Motivating Example Using Ecological Momentary Assessment (EMA) Data

DFA models with random effects are often used to represent intraindividual dynamics and interindividual differences in intraindividual dynamics in ecological momentary assessment (EMA) studies in which participants are asked to provide ratings of their subjective states “in the moment” and in naturalistic settings. As part of a larger study (see Diener et al., 1995), 174 college students (93 male and 81 female; average age = 20.24 years, SD = 1.41) were asked to provide daily EMA ratings over 52 days. The participants were asked to indicate how often they had experienced a series of emotions since the last assessment using a 7-point Likert scale (with 1 = “None” to 7 = “Always”). Their responses on 8 ordinal items, namely, the items *joy*, *contentment*, *love*, *affection*, *unhappiness*, *anger*, *depression*, and *anxiety*, were used in the present application. The first four items were items that served to indicate an underlying unobserved (i.e., latent) variable, positive emotion (PE), while the last four items were as indicators of individuals' latent negative emotion (denoted below as NE). The proportion of missing data in the sample was about 11.2%.

Chow et al. (2011) used a nonlinear DFA model inspired by a theoretical model of affects proposed by Zautra and colleagues (Zautra et al., 2000) to represent the interrelations between the PE and NE factors in these data. The proposed model postulates that the relative separation between positive emotion (PE) and negative emotion (NE) changes dynamically as a function of the previous levels of the emotion of opposite valence as:

$$\begin{aligned} PE_{it} &= \left[b_{11i} + b_{12i} \frac{\exp(NE_{i,t-1})}{1 + \exp(NE_{i,t-1})} \right] PE_{i,t-1} + \zeta_{1it}, \\ NE_{it} &= \left[b_{22i} + b_{21i} \frac{\exp(PE_{i,t-1})}{1 + \exp(PE_{i,t-1})} \right] NE_{i,t-1} + \zeta_{2it}, \\ \zeta_{it} &= (\zeta_{1it}, \zeta_{2it})^T \sim N(\mathbf{0}, \mathbf{\Psi}_{\zeta}), \mathbf{\Psi}_{\zeta} = \begin{bmatrix} \psi_{\zeta 11} & \psi_{\zeta 12} \\ \psi_{\zeta 12} & \psi_{\zeta 22} \end{bmatrix}, \end{aligned} \quad (1)$$

where PE_{it} and NE_{it} represent individual i 's unobserved positive and negative emotion factors at time t , respectively; ζ_{1it} and ζ_{2it} are process noise terms. Moreover, b_{11i} and b_{22i} are the baseline first-order autoregressive (AR(1)) parameters of individual i , whereas b_{12i}

and b_{21i} are deviations in AR(1) parameters when individual i manifested high levels of $NE_{i,t-1}$ and $PE_{i,t-1}$ at time $t-1$, respectively. This dynamic model suggests that if an individual's PE (or NE) was high at time $t-1$, the high PE (NE) affects the subsequent dynamics of the individual's NE (PE) by altering the AR(1) parameter of the latter. That is, we postulate that the high NE that arises under high-stress situations would, in turn, change the dynamics of PE at the next time point by altering the value of PE's AR(1) parameter by a magnitude of b_{12i} . Whether a reciprocal effect from PE to NE also exists at very high values of PE is reflected in the magnitude of b_{21i} .

The parameters b_{11i} , b_{12i} , b_{21i} , and b_{22i} are time series parameters that are often constrained to satisfy the stability conditions (Lütkepohl, 2005), resulting in complex restrictions of range in these parameters' possible values. The stability constraints ensure that the processes under model show statistical properties such as means and covariance functions that are finite in values and invariant over time, among other characteristics. In the absence of such constraints, the processes of interest would show increasing variances over time and possibly amplify to infinite values — an unlikely characteristic for the kinds of processes targeted by EMA studies that span weeks. As a result, the distributions of the dynamic parameters may not be normal or even symmetric.

Almost all of the applications utilizing DFA models, including the study by Chow et al. (2011), involved intensive longitudinal data that span weeks. The data were typically assumed to conform to ignorable missingness mechanisms in these studies. In other words, the data were assumed to be missing completely at random, or missing at random (MAR) depending only on observed information in the data (Little and Rubin, 2002). Myriad reasons abound for why nonignorable missingness may arise in an intensive longitudinal context. For instance, when individuals are asked to provide repeated ratings on their current emotions over weeks, it is plausible, perhaps even probable, for individuals to opt not to respond when they are feeling too bad (e.g., after the death of a friend), or too good (e.g., while away on a vacation). Even if the model can be extended to allow for nonignorable missingness, it is difficult to falsify the hypothesized model of non-ignorable missingness due to the lack of sufficient information from the data (e.g., Ibrahim et al., 1999; Shi et al., 2009). Conducting sensitivity analysis provides critical information on the tenability of such assumptions in the models.

Existing methods for performing sensitivity analysis on missing data models rely either on analytic sensitivity measures that quantify changes in some localized aspects of the model (e.g., biases or changes in selected parameters; Rotnitzky and Wypij, 1994), use Monte Carlo simulations to perform sensitivity analysis (Kaplan and George, 1995; De Beuckelaer and Swinnen, 2011), or compare inferential results under different missing data models (Blozis et al., 2013). Thus, most of these approaches are computationally expensive (e.g., involve the fitting of multiple models or Monte Carlo replications of the same model); or are limited in using very localized measures (e.g., univariate changes in selected modeling parameters) obtained by perturbing a limited number of modeling components — usually in a sequential manner. The present article is the first adaptation of the local influence approach by Zhu et al. (2011) to perform simultaneous sensitivity analysis of multiple modeling components in DFA models under the DP prior and with both continuous and ordinal data.

The rest of this paper is organized as follows. We first present the general modeling framework within which the motivating model of emotions can be viewed as a special case. We then introduce the general perturbation method by Zhu et al. (2011) to highlight how this method can be used for local sensitivity analysis in DFA models and beyond. This is followed by five illustrations in which the local influence method is used to diagnose the presence of outliers and sources of model misspecification. Finally, we present results from our motivating application and some concluding remarks.

Nonlinear DFA Model with Random Effects and Dirichlet Process (DP) Prior

A DFA model is composed of two submodels: a measurement model that specifies how a vector of latent variables – in particular, factors, denoted as η_{it} for person i at time t – are apprehended through a vector of manifest indicators y_{it} and a dynamic (“structural”) model that portrays the over-time interrelations among the latent factors. The measurement equation linking the latent variables to their manifest indicators is expressed as

$$y_{it} = A c_{it} + \Lambda \eta_{it} + \epsilon_{it}, \quad \epsilon_{it} \stackrel{\text{i.i.d.}}{\sim} N(\mathbf{0}, \Psi_\epsilon), \quad t = 1, \dots, T_i, \quad i = 1, \dots, n, \quad (2)$$

where T_i is the number of occasions measured for the i th subject, n is the number of subjects, y_{it} is a $p \times 1$ vector of manifest variables measured at time t for the i th subject, c_{it} is an $s \times 1$ vector of covariates which may come from a continuous or discrete distribution, and η_{it} is a $q \times 1$ vector of latent variables of interest. Moreover, A is a $p \times s$ matrix of unknown parameters, Λ is a $p \times q$ matrix of factor loadings, and $\Psi = \text{diag}(\psi_{e1}, \dots, \psi_{ep})$ is a $p \times p$ covariance matrix of errors ϵ_{it} . It is assumed that η_{it} and ϵ_{it} are independent. The covariates can be explanatory or other kinds of variables that are helpful in achieving a better model for relating y_{it} to η_{it} .

To accommodate both continuous and ordered categorical data, let $y_{it} = (y_{(1)it}^T, y_{(2)it}^T)^T$, where $y_{(1)it} = (y_{it1}, \dots, y_{itk_{it}})^T$ is a subset of continuous indicators that cannot be directly observed, whilst $y_{(2)it}$ can be directly observed. For $j = 1, \dots, k_{it} - p$, the unobserved continuous latent variable, y_{itj} , is linked to the observed ordinal indicator, u_{itj} , characterized by κ_j categories, as $u_{itj} = m$ if $\tau_{j,m-1} < y_{itj} \leq \tau_{j,m}$. τ is a vector of threshold parameters for all the ordinal items, held invariant across subjects; and we define $U_{it} = (u_{it1}, \dots, u_{itk_{it}})^T$ as a vector that contains person i 's k_{it} ordinal indicators at time t .

The dynamic model in a DFA describes the interrelations among the latent factors as:

$$\eta_{it} = f_t(\eta_{i,t-1}, b_i, \theta_\eta) + \zeta_{it}, \quad \zeta_{it} \stackrel{\text{i.i.d.}}{\sim} N_q(\mathbf{0}, \Psi_\zeta), \quad (3)$$

where b_i is a $n_b \times 1$ random vector of subject-specific parameters modeled using the DP; ζ_{it} is a $q \times 1$ vector of process noise components; $f_t(\cdot)$ is a vector of time-varying, differentiable linear or nonlinear functions describing the latent variables at time t in terms of their

previous levels at time $t - 1$. Moreover, θ_η is a vector of parameters that is held invariant over time and subjects, and Ψ_ζ is a $q \times q$ covariance matrix of residuals/process noise ζ_{it} . It is assumed that $\eta_{i,t-1}$ is independent of b_i , and ζ_{it} is independent of $\eta_{i,t-1}$ and b_i . The model of emotions shown in Equation (1), in which $\eta_{it} = (PE_{it}, NE_{it})^T$, is a special case of Equation (3).

DP prior for the time series parameters

In applications utilizing a DFA model, the key inferential questions of interest typically revolve around the person-specific time series parameters in $b_i = (b_{11i}, b_{22i}, b_{12i}, b_{21i})^T$. Adding to the challenge, however, are these parameters' restrictions in range, and possible roles in biasing key interpretations concerning interindividual differences in dynamics under violations of distributional assumptions. To help “robustify” (Antoniak, 1974; Ferguson, 1973) and assess the appropriateness of common parametric assumptions (Dunson, 2008; Karabotsos and Walker, 2013) on b_i , we use a Dirichlet process (DP) prior for b_i . That is, we let $b_i \sim \text{DP}(\alpha P_0)$, where P_0 is a base distribution (e.g., a multivariate normal distribution) that serves as a starting point for constructing a nonparametric distribution for b_i . Parameter α (with $\alpha > 0$) represents the weight that a researcher assigns a priori to the base distribution, and it reflects the researcher's certainty of P_0 as the distribution of b_i .

The way in which the DP prior serves as a nonparametric prior can be understood in terms of Sethuraman's (1994) stick-breaking representation. The stick-breaking representation provides an alternative way of conceiving $P \sim \text{DP}(\alpha P_0)$ as

$$\mathcal{P}(\cdot) = \sum_{g=1}^{\infty} \pi_g \delta_{Z_g}(\cdot), \quad \text{with } Z_g \stackrel{\text{i.i.d.}}{\sim} P_0, \quad (4)$$

where $\delta_{Z_g}(\cdot)$ denotes a unit point mass at Z_g , Z_g is a $n_b \times 1$ vector consisting of possible values of b_i that are assigned to the g th latent cluster (i.e., stick), and π_g is a random probability weight between 0 and 1. Equation (4) thus posits that $\mathcal{P}$ consists of a series of point masses—or “sticks”—of various lengths concentrated at different values of Z_g .

For empirical estimation purposes, a truncated Dirichlet processes may be used to approximate Equation (4) as (Ishwaran and Zarepour, 2000; Ishwaran and James, 2001; Lee et al., 2007):

$$\mathcal{P}(\cdot) \approx \mathcal{P}_G(\cdot) = \sum_{g=1}^G \pi_g \delta_{Z_g}(\cdot), \quad 1 \leq G < \infty, \quad (5)$$

where π_g is obtained as

$$\pi_1 = v_1, \quad \pi_g = v_g \prod_{l < g} (1 - v_l) \text{ for } g = 2, \dots, G \text{ with } \quad (6)$$

$$v_g \stackrel{\text{i.i.d.}}{\sim} \text{Beta}(1, \alpha), \text{ for } g = 1, \dots, G-1 \text{ and } \quad (7)$$

$v_G = 1$ so that $\sum_{g=1}^G \pi_g = 1$. In the simple case where $G = 2$, $v_2 = 1$, v_1 is a random weight between 0 and 1 drawn from the Beta distribution; π_1 , the proportion or length of the first stick, is equal to v_1 and by Equation (6), $\pi_2 = v_2(1 - v_1) = 1 - v_1$. That is, v_1 is the proportion of a unit probability stick that is broken off and assigned to Z_1 , and $1 - v_1$ is the remaining of the stick that is assigned to Z_2 . In practice, the value of G is either set to a large, predetermined value (e.g., $G = 120$), or chosen empirically. Our previous experience showed that a value of $G = 180$ is more than adequate for the model considered in the present context. Following the idea behind data augmentation (Tanner and Wong, 1987), a categorical latent variable, L_i , is included in the model to indicate the latent cluster (or class) membership of $\mathbf{b}_i$ (Chow et al., 2011; Ishwaran and James, 2001; Lee et al., 2007). The values of the random effects in $\mathbf{b}_i$ are given by $\mathbf{b}_i = \mathbf{Z}_{L_i}$ and can be deduced once the vector's cluster membership is known. Further, $L_i | \pi \stackrel{\text{i.i.d.}}{\sim} \sum_{g=1}^G \pi_g \delta_g(\cdot)$.

Through different choices of the base distribution, P_0 , and allocations of the weights, π_g , a variety of distributional forms can be approximated by the discrete probability measure $\mathcal{P}$. The allocation of π_g is, in turn, governed by the random weight, α . Typically, a gamma hyperprior is specified for α and the associated hyperparameters for this gamma prior reflect a researcher's best guess on how π_g should be distributed, and consequently, how the point masses in Equation (5) should resemble the base distribution, P_0 (for further details see Dunson (2008) and Chow et al. (2011)).

Missing data assumptions

To deal with missing data, we partition the manifest continuous and categorical responses to portions corresponding to observed and missing data. Specifically, we let $\mathbf{y}_{(2)it} = \{\mathbf{y}_{(2)it, \text{obs}}, \mathbf{y}_{(2)it, \text{mis}}\}$, where $\mathbf{y}_{(2)it, \text{obs}}$ and $\mathbf{y}_{(2)it, \text{mis}}$ are the observed and missing continuous manifest variables, respectively. Similarly we partition $\mathbf{U}_{it}$ into $\{\mathbf{U}_{it, \text{obs}}, \mathbf{U}_{it, \text{mis}}\}$, where $\mathbf{U}_{it, \text{obs}}$ and $\mathbf{U}_{it, \text{mis}}$ are the observed and missing binary/ordinal data in $\mathbf{U}_{it}$, respectively, and $\mathbf{y}_{(1)it, \text{obs}}$ and $\mathbf{y}_{(1)it, \text{mis}}$ are subvectors of $\mathbf{y}_{(1)it}$ containing the unobserved continuous counterparts to $\mathbf{U}_{it, \text{obs}}$ and $\mathbf{U}_{it, \text{mis}}$, respectively.

Let $\mathbf{r}_{it} = (r_{it1}, \dots, r_{itp})^T$ be a vector of observed missing status indicators for $\mathbf{U}_{it}$ and $\mathbf{y}_{(2)it}$ such that $r_{itj} = 1$ if u_{itj} (or y_{itj}) is missing and 0 otherwise for $j = 1, \dots, k_{it}$ (or $k_{it} + 1, \dots, p$). The missing data mechanism is assumed to conform to the density $p(\mathbf{r}_{it} | \mathbf{y}_{(1)it, \text{mis}}, \mathbf{y}_{(2)it, \text{obs}}, \mathbf{U}_{it, \text{obs}}, \boldsymbol{\phi})$, in which $\boldsymbol{\phi}$ is a parameter vector. We consider a nonignorable missing data mechanism wherein $p(\mathbf{r}_{it} | \mathbf{y}_{(1)it, \text{mis}}, \mathbf{y}_{(2)it, \text{obs}}, \mathbf{U}_{it, \text{obs}}, \boldsymbol{\phi})$ is specified to be a multinomial distribution with 2^p cell probabilities, written as

$$p(\mathbf{r}_{it} | \mathbf{y}_{(1)it, \text{mis}}, \mathbf{y}_{(2)it}, \mathbf{U}_{it, \text{obs}}, \boldsymbol{\varphi}) = \prod_{j=1}^p \{ \text{pr}(r_{itj} = 1 | \mathbf{y}_{(1)it, \text{mis}}, \mathbf{y}_{(2)it}, \mathbf{U}_{it, \text{obs}}, \boldsymbol{\varphi}) \}^{r_{itj}} \\ \times \{ 1 - \text{pr}(r_{itj} = 1 | \mathbf{y}_{(1)it, \text{mis}}, \mathbf{y}_{(2)it}, \mathbf{U}_{it, \text{obs}}, \boldsymbol{\varphi}) \}^{1 - r_{itj}},$$

where $\text{pr}(r_{itj} = 1 | \mathbf{y}_{(1)it, \text{mis}}, \mathbf{y}_{(2)it}, \mathbf{U}_{it, \text{obs}}, \boldsymbol{\varphi})$ is modeled by

$$\text{logit}\{\text{pr}(r_{itj} = 1 | \mathbf{y}_{(1)it, \text{mis}}, \mathbf{y}_{(2)it}, \mathbf{U}_{it, \text{obs}}, \boldsymbol{\varphi})\} = \boldsymbol{\varphi}_1^T \mathbf{x}_{it} + \boldsymbol{\varphi}_2^T \mathbf{x}_{i, t-1}, \quad j = 1, \dots, p. \quad (8)$$

in which $\mathbf{x}_{it}$ includes selected elements in $\mathbf{y}_{(1)it, \text{mis}}, \mathbf{y}_{(2)it}$ and $\mathbf{U}_{it, \text{obs}}$, and $\mathbf{x}_{i, t-1}$ includes selected elements in $\mathbf{y}_{(1)i, t-1, \text{mis}}, \mathbf{y}_{(2)i, t-1}$, and $\mathbf{U}_{i, t-1, \text{obs}}$; $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_1^T, \boldsymbol{\varphi}_2^T)^T$ is a $d \times 1$ vector of regression coefficients. Here, the parameter vector $\boldsymbol{\varphi}$ is used to capture the dependencies of an individual's missing data status on the observed and unobserved indicators. In addition, we assume that the distribution of $\mathbf{r}_{it}$ does not depend on the latent variables in $\boldsymbol{\eta}_{it}$ and the random effects variables in $\mathbf{b}_i$ (e.g., Lee and Tang, 2006). To limit the complexity of the model, we assume no missingness in the covariates in $\mathbf{c}_{it}$ (see Equation 2).

Gibbs sampling procedures

Given that the focus of this paper on local influence analysis as opposed to estimation procedures for the proposed model, we only provide a brief overview of the Gibbs sampling procedures used for model estimation purposes. Further estimation details can be found elsewhere (Chow et al., view).

We define $\mathbf{Z} = (\mathbf{Z}_1, \dots, \mathbf{Z}_G)$ as an $n_b \times G$ matrix containing values of $\mathbf{b}_i$, and the vectors, $\boldsymbol{\pi} = (\boldsymbol{\pi}_1, \dots, \boldsymbol{\pi}_G)$, and $\mathbf{L} = (\mathbf{L}_1, \dots, \mathbf{L}_n)$. Further, we define the following arrays, as containing all corresponding elements for $i = 1, \dots, n$, and (when applicable) $t = 1, \dots, T$: $\mathbf{H} = \{\boldsymbol{\eta}_{it}\}$, $\mathbf{b} = \{\mathbf{b}_i\}$, $\mathbf{U}_{\text{obs}} = \{\mathbf{U}_{it, \text{obs}}\}$, $\mathbf{Y}_{(1)\text{obs}} = \{\mathbf{y}_{(1)it, \text{obs}}\}$, $\mathbf{Y}_{(2)\text{obs}} = \{\mathbf{y}_{(2)it, \text{obs}}\}$, $\mathbf{r} = \{\mathbf{r}_{it}\}$, $\mathbf{C} = \{\mathbf{c}_{it}\}$, $\mathbf{Y} = \{\mathbf{y}_{it}\}$, and $\mathbf{Y}_{\text{mis}} = \{\mathbf{y}_{(1)it, \text{mis}}, \mathbf{y}_{(2)it, \text{mis}}\}$. The unknown parameters are $\boldsymbol{\theta} = \{\boldsymbol{\theta}_y, \boldsymbol{\theta}_\eta, \boldsymbol{\Psi}_\zeta, \boldsymbol{\theta}_b\}$, in which $\boldsymbol{\theta}_y$ contains all unknown parameters in $\mathbf{A}$, $\boldsymbol{\Lambda}$, $\boldsymbol{\tau}$ and $\boldsymbol{\Psi}_\epsilon$; $\boldsymbol{\theta}_b$ contains all unknown parameters relating to the distribution of $\mathbf{b}_i$, and $\boldsymbol{\theta}_\eta$ and $\boldsymbol{\Psi}_\zeta$ are as defined earlier.

Inspired by the idea of data augmentation (Tanner and Wong, 1987), we treat $\mathbf{H}$, $\mathbf{L}$, $\mathbf{Z}$ and $\mathbf{Y}_{(1)\text{obs}}$ as hypothetical missing data and use the Gibbs sampler to generate samples from the joint posterior distribution by sequentially sampling from a series of conditional distributions. Each of these conditional distributions, as detailed in Supplementary Appendix A, features the distribution of one unknown modeling component (e.g., $\mathbf{H}$) or a block of highly correlated modeling components (e.g., some of the interdependent components that appear in the DP prior) conditional on all other pertinent unknown modeling components and the observed data. In cases where the conditional distributions involved do not conform to standard distributional forms, the Metropolis–Hasting (MH) algorithm is used to draw samples from appropriate proposal distributions. All prior distributions are also described in

Supplementary Appendix A; their hyperparameters are summarized in Supplementary Appendix B.

After K_0 burn-in samples, the empirical distribution of the remaining K_1 samples $\{\mu_Z^{(\kappa)}, \Psi_Z^{(\kappa)}, \alpha^{(\kappa)}, \pi^{(\kappa)}, Z^{(\kappa)}, L^{(\kappa)}, H^{(\kappa)}, \theta^{(\kappa)}, Y_{\text{mis}}^{(\kappa)}, \tau^{(\kappa)}, \phi^{(\kappa)}; K_0 < \kappa \leq K_0 + K_1\}$ is taken as an approximation of the joint posterior distribution, on which Bayesian estimates and inference are based. For instance, sample means and percentiles are used as approximation to posterior means and credible intervals, and other summary statistics computed using these samples are used to replace integration and other procedures that are otherwise analytically intractable. Throughout, estimated potential scale reduction (EPSR) values (Gelman and Rubin, 1992) across three chains is used to monitor the convergence of the Gibbs algorithm presented above, with values close to 1.1 taken as a sign of relative stability of the estimates.

Bayesian Local Influence Analysis

The Bayesian local influence method developed by Zhu et al. (2011) enables researchers to perform sensitivity analysis by incurring simultaneous changes (i.e., perturbations) to the data and multiple elements of a model – including prior and sampling distributions – all within a single fitting of the Bayesian model of choice. To apply this method to DFA models, we first augment the observed data set, $D_{\text{obs}} = \{U_{\text{obs}}, Y_{(2)\text{obs}}, r, C\}$ with “missing” data components in $D_{\text{mis}} = \{H, L, Z, Y_{(1)\text{obs}}, \text{ and } Y_{\text{mis}}\}$ to obtain the complete data set $D_{\text{com}} = \{H, L, Z, U_{\text{obs}}, Y, r, C\}$. All the parameters are collected into $\boldsymbol{\vartheta} = \{\theta_y, \theta_r, \Psi_\zeta, \tau, \phi, \alpha, \mu_Z, \Psi_Z, \pi\}$.

Consider a perturbation vector, $\omega = (\omega_D^T, \omega_P^T, \omega_S^T)^T$, with subvectors containing, respectively, perturbations to individual scores in D_{com} , perturbations to the priors of unknown parameters in $\boldsymbol{\vartheta}$, and perturbations to the sampling distributions. The corresponding perturbed joint density function can be written as

$$p(D_{\text{com}}, \boldsymbol{\vartheta} | \omega) = p(\boldsymbol{\vartheta} | \omega_P) \prod_{i=1}^n \prod_{t=1}^{T_i} p(D_{it, \text{com}} | \boldsymbol{\vartheta}, \omega_D, \omega_S), \quad (9)$$

wherein $p(D_{\text{com}}, \boldsymbol{\vartheta} | \omega)$ essentially spells out a model that describes the joint density of D_{com} and $\boldsymbol{\vartheta}$ under the perturbation scheme in ω , and it has to integrate to 1 under that perturbation scheme, namely, $\int p(D_{\text{com}}, \boldsymbol{\vartheta} | \omega) dD_{\text{com}} d\boldsymbol{\vartheta} = 1$. Moreover, ω^0 , consisting of ω_D^0 , ω_P^0 , and ω_S , represents no perturbation to any modeling element. For the purpose of local influence analysis, we focus on $\log p(D_{\text{com}}, \boldsymbol{\vartheta} | \omega)$, the log of the perturbed joint density.

We consider both single-case and global perturbation schemes in our application. Single-case perturbation scheme refers to any scheme that independently perturbs individual observations (Verbeke et al., 2001), typically to identify influential observations (Shi et al., 2009). Examples include casewise perturbations to individual subjects' observed scores, latent variable scores, and random effects. In this case, it is necessary to define an

individual-specific perturbation vector, ω_{Di} , for the i th subject such that $\omega_D = \{\omega_{Di} : i = 1, \dots, n\}$.

A global perturbation scheme refers to any scheme that perturbs all observations simultaneously to assess the robustness of model assumptions to minor perturbations (Troxel et al., 2004; Copas and Eguchi, 2005; Shi et al., 2009), administered via ω_S and/or ω_P . Examples of ω_S may include perturbations of the assumption of MAR in the direction of not missing at random (NMAR; Little and Rubin, 2002). Finally, ω_P may contain perturbations to the hyperparameters and functional forms of the prior distributions implicated in the Bayesian analysis.

Consider a perturbation model, $\mathcal{M} \triangleq \{p(D_{\text{com}}, \boldsymbol{\theta} | \boldsymbol{\omega}) : \boldsymbol{\omega} \in \mathbb{R}^{\mathcal{T}}\}$, that imposes $\mathcal{T}$ simultaneous perturbations on the joint density of D_{com} and $\boldsymbol{\theta}$. In this context, $\mathcal{M}$ can be regarded as a $\mathcal{T}$ -dimensional manifold. The $\mathcal{T}$ perturbations in $\boldsymbol{\omega}$ result in $\mathcal{T}$ tangent functions on $\mathcal{M}$ (Amari, 1990) along each of the specified perturbations, namely, $\partial_{\omega_j} \log p(D_{\text{com}}, \boldsymbol{\theta} | \boldsymbol{\omega})$, where ∂_{ω_j} is an abbreviated expression used throughout to denote $\partial / \partial \omega_j$. These tangent functions quantify local changes to the joint density of D_{com} and $\boldsymbol{\theta}$ with respect to each perturbation element. The corresponding vector provides a starting point for conducting a sensitivity analysis but these tangent functions may have arbitrarily different metrics along each of the $\mathcal{T}$ directions; and they may not be orthogonal to each other. That is, imposing perturbation on a single element may inadvertently lead to perturbation(s) on one or more other elements.

To ensure the orthogonality of the perturbation model along all $\mathcal{T}$ directions, it is helpful to inspect the $\mathcal{T}^2$ quantities,

$$\begin{aligned} g_{jk}(\boldsymbol{\omega}) &= E_{\boldsymbol{\omega}}[\partial_{\omega_j} \log p(D_{\text{com}}, \boldsymbol{\theta} | \boldsymbol{\omega}) \partial_{\omega_k} \log p(D_{\text{com}}, \boldsymbol{\theta} | \boldsymbol{\omega})] \\ &= -E_{\boldsymbol{\omega}}[\partial_{\omega_j}^2 \log p(D_{\text{com}}, \boldsymbol{\theta} | \boldsymbol{\omega})], \text{ for } j, k = 1, \dots, \mathcal{T}, \end{aligned} \quad (10)$$

which form the so-called *metric tensor* of $\mathcal{M}$ (Amari, 1990), where $E_{\boldsymbol{\omega}}$ denotes expectation taken with respect to $p(D_{\text{com}}, \boldsymbol{\theta} | \boldsymbol{\omega})$. The $\mathcal{T} \times \mathcal{T}$ matrix $G(\boldsymbol{\omega}) = (g_{jk}(\boldsymbol{\omega}))$, with its (j, k) element given by (10), is simply the Fisher information matrix with respect to $\boldsymbol{\omega}$. The element $g_{jk}(\boldsymbol{\omega})$ reflects the association between ω_j and ω_k (Zhu et al., 2011). The metric tensor essentially summarizes the curvature, or uncertainty in the perturbed joint density with respect to each perturbed element in $\boldsymbol{\omega}$. To aid interpretation, it is helpful to specify a perturbation model such that $G(\boldsymbol{\omega}^0)$ equals $cI_{\mathcal{T}}$, where c is an arbitrary constant. In other words, it is helpful to transform the perturbations such that the $\mathcal{T}$ perturbations that are to be administered are orthogonal to each other, and that the same amount of perturbation is administered in all $\mathcal{T}$ directions. In practice, this is not always possible, but this can be accomplished by revising the perturbation scheme to yield a new perturbation vector, $\tilde{\boldsymbol{\omega}}$, defined as

$$\tilde{\omega}(\omega) = \omega^0 + G(\omega^0)^{1/2}(\omega - \omega^0), \quad (11)$$

such that $G(\tilde{\omega})$ evaluated at ω^0 equals $I_{\mathcal{T}}$ (Zhu et al., 2007, 2011). That is, $\tilde{\omega}(\omega)$ is an appropriate perturbation at least at $\tilde{\omega}(\omega^0) = \omega^0$.

Let $\mathbf{h}$ be a $\mathcal{T} \times 1$ vector of binary constants distinguishing elements in the model that are perturbed ($= 1$) and those that are not ($= 0$). Consider $\omega(a)$, which applies perturbations of magnitude a to the perturbation vector, ω , in the directions as dictated in $\mathbf{h}$, with $\omega(0) = \omega^0$. Next, consider possible choices of local influence measures that can help quantify the impacts of $\omega(a)$. This requires that a researcher selects some objective functions of interest to summarize the results of the sensitivity analysis. The objective functions may be a model comparison criterion such as the Bayes factor, the posterior means of particular parameters of interest, or a function to assess the divergence in posterior distributions with and without perturbations, just to name a few examples. Using an objective function such as the Bayes factor provides a way to assess the impacts of $\omega(a)$ on overall model fit, whereas objective functions such as the posterior means of selected model parameters provide a glimpse into more localized consequences of the perturbations.

Let $f(\omega(a))$ denote an objective function of choice used to quantify the effects of the perturbation and let $\nabla f \triangleq \nabla_{\omega} f(\omega^0)$ be the score vector of the objective function with respect to ω at ω^0 . Using Taylor series expansion and retaining only the linear term, $f(\omega(a))$ along the directions as dictated in $\mathbf{h}$ may be approximated as $f(\omega^0) + a \nabla_f^T \mathbf{h}$. Consequently, the overall local change in the objective function with respect to ω around ω^0 is then summarized in $\nabla_f^T \mathbf{h}$. Capitalizing on $\nabla_f^T \mathbf{h}$, one may construct a “standardized” *first-order influence measure* (FI) that takes into consideration the uncertainties contained in $G(\omega^0)$ along the directions dictated in $\mathbf{h}$ to yield

$$FI_{f, \mathbf{h}} = FI_{f(\omega^0), \mathbf{h}} = \frac{\mathbf{h}^T \nabla_f W_f \nabla_f^T \mathbf{h}}{\mathbf{h}^T G \mathbf{h}}, \quad (12)$$

where $G = G(\omega^0)$ and W_f is a positive semi-definite scaling matrix. For the perturbation $\tilde{\omega}(\omega)$ in (11), $FI_{f, \mathbf{h}}$ takes on the form of:

$$FI_{f(\tilde{\omega}), \mathbf{h}}|_{\tilde{\omega} = \omega^0} = \frac{\mathbf{h}^T G^{-1/2} \nabla_f W_f \nabla_f^T G^{-1/2} \mathbf{h}}{\mathbf{h}^T \mathbf{h}}. \quad (13)$$

In cases where the perturbation scheme involves a large number of simultaneous perturbations, one may inspect $FI_{f, \mathbf{e}_j}$, where $\mathbf{h} = \mathbf{e}_j$ is set to be a $\mathcal{T} \times 1$ vector with the j^{th}

component set to 1 and 0 elsewhere (Zhu et al., 2007, 2011). In this way, the effect attributable to each of the perturbed components can be quantified.

When $\nabla_f = 0$, we can use a *second-order influence measure* (SI) that quantifies the “standardized” *local curvature* in the objective function with respect to ω around ω^0 as:

$$SI_{f, \mathbf{h}} = SI_{f(\omega^0), \mathbf{h}} = \frac{\mathbf{h}^T \mathbf{H}_f \mathbf{h}}{\mathbf{h}^T \mathbf{G} \mathbf{h}}, \quad (14)$$

where $\mathbf{H}_f = \partial_{\omega}^2 f(\omega^0)$. Particularly, for the appropriate perturbation $\tilde{\omega}(\omega)$ in (11), $SI_{f, \mathbf{h}}$ assumes the form of

$$SI_{f(\tilde{\omega}), \mathbf{h}}|_{\tilde{\omega} = \omega^0} = \frac{\mathbf{h}^T \mathbf{G}^{-1/2} \mathbf{H}_f \mathbf{G}^{-1/2} \mathbf{h}}{\mathbf{h}^T \mathbf{h}}. \quad (15)$$

Similar to first-order influence measures, second-order influence measures summarize the curvatures of the influence measures when particular elements in the model are perturbed. Larger values on a second-order influence measure correspond to greater (steeper) curvatures in the objective influence function of choice around the neighborhood of ω^0 and thus, larger influence or more “precipitous” changes in the objective function under perturbations. Similar to the first-order influence measure, we can examine SI_{f, e_j} to quantify the effect of each of the perturbed component(s) in turn.

To compute the local influence measures described above, the key computations lie in computing $\mathbf{G}$, ∇_f , and/or $\mathbf{H}_f$. In Appendix A, we describe how to compute ∇_f and/or $\mathbf{H}_f$ for local influence measures based on three objective functions, namely, the Bayes factor, the posterior means of any function, $d(\boldsymbol{\theta})$, of $\boldsymbol{\theta}$ (e.g., square-root of variance parameters), and the ϕ -divergence of the posterior distributions with perturbation relative to the posterior distributions under no perturbation (Zhu et al., 2011). All of these local influence measures are approximated using summary statistics computed using the MCMC samples obtained during the model estimation after the burn-in iterations.

Illustrative Examples

To elucidate properties of our Bayesian local influence approach, we begin with a simple “toy” example, followed by four additional illustrative examples based on perturbing components of the DFA model shown in Equations (1)–(2).

Illustration I: Coin Toss Example

Consider a coin that is tossed N times in two experiments, each time with the same success probability, θ , and the number of successes are recorded as z_1 and z_2 . The likelihood function of z_1 and z_2 is given by

$$p(z_1, z_2 | \theta) = \frac{\theta^{z_1} (1-\theta)^{N-z_1} \theta^{z_2} (1-\theta)^{N-z_2}}{(N+1)^2 B(z_1+1, N-z_1+1) B(z_2+1, N-z_2+1)},$$

where $B(a, b)$ is the Beta function with parameters a and b . To make Bayesian inference on θ , suppose we consider the beta prior, $p(\theta) \sim \text{Beta}(0, 1)$, for θ . This is a very informative prior with θ showing a point mass probability concentrating at the value of 0.

To evaluate how sensitive our conclusions are to a possible change in success probability from the first to the second round, we consider perturbation to the success probability in the second round. In this case, the corresponding perturbed likelihood function is given by

$$p(z_1, z_2 | \omega) = \frac{\theta^{z_1} (1-\theta)^{N-z_1} (\omega\theta)^{z_2} (1-\omega\theta)^{N-z_2}}{(N+1)^2 B(z_1+1, N-z_1+1) B(z_2+1, N-z_2+1)}.$$

Here $\omega^0 = 1$ represents no perturbation. The first- and second-order partial derivatives with respect to ω are given by

$$\begin{aligned} \frac{\partial \log p(z_1, z_2 | \omega)}{\partial \omega} &= \frac{z_2}{\omega} - \frac{\theta(N-z_2)}{1-\omega\theta}, \quad \frac{\partial \log p(z_1, z_2 | \omega)}{\partial \omega} \Big|_{\omega=1} = z_2 - \frac{\theta(N-z_2)}{1-\theta} = N \frac{z_2 - \theta}{1-\theta}, \\ \frac{\partial^2 \log p(z_1, z_2 | \omega)}{\partial \omega^2} &= -\frac{z_2}{\omega^2} - \frac{\theta^2(N-z_2)}{(1-\omega\theta)^2}, \quad \frac{\partial^2 \log p(z_1, z_2 | \omega)}{\partial \omega^2} \Big|_{\omega=1} = -z_2 - \frac{\theta^2(N-z_2)}{(1-\theta)^2}. \end{aligned}$$

Thus, it follows from $E(z_2) = N\theta$ that

$$g(\omega^0) = E_{\theta, z_2} \left(z_2 + \frac{\theta^2(N-z_2)}{(1-\theta)^2} \right) = E_{\theta} \left(E(z_2) + \frac{\theta^2(N-E(z_2))}{(1-\theta)^2} \right) = NE_{\theta} \left(\frac{\theta}{1-\theta} \right),$$

where E_{θ, z_2} and E_{θ} represent the expectations taken with respect to the joint probability

density $p(\theta, z_2)$ and the prior, $p(\theta)$, respectively. In this case, the value of $g(\omega^0) = 0$ indicates that there is little curvature or uncertainty in the perturbed likelihood function in the presence of such strong prior information. Substituting $G(\omega^0) = g(\omega^0) = 0$ into the calculation of Equation 12 or 13 yields a non-finite value. Thus, in this simple case, there is nothing much to be gained from the local sensitivity analysis because in the presence of such a strong prior, our conclusions are not sensitive to a change in success probability from the first to the second round. In contrast, if we assume a beta prior with parameters α and β for θ , i.e., $p(\theta) \sim \text{Beta}(\alpha, \beta)$, it can be shown that $g(\omega^0) = N\alpha/(\beta-1)$. In this case, the curvature depends on the selection of hyperparameters in the beta prior. A slight extension to this example, demonstrating the sensitivity of inferential results to perturbations to the beta prior when uninformative hyperparameters are used, is provided in Supplementary Appendix C, with additional matlab code included for didactic purposes.

Illustration II: Cross-Loading Detection in Confirmatory Factor Analysis with Ordinal Items

The DFA model considered in this study includes a confirmatory factor analysis (CFA) model with ordinal items as its measurement model. CFA is a popular statistical technique for multivariate data analysis, and it starts with strong, theoretically informed assumptions on the structure of the factor loading matrix, Λ . In many CFA applications, each latent factor is constrained to only load on one manifest item, with all cross-loadings set to zero. Recently, there has been a surge of interest within the Bayesian CFA context to consider more flexible approaches for detecting misspecification in factor loading structures, such as cross-loadings fixed at 0 that, when freed up, may improve model fit and reduce spurious correlations among factors (Lu et al., 2016; Muthén and Asparouhov, 2012).

To demonstrate how the proposed local influence approach may be used to detect cross-loadings that would otherwise be falsely constrained to be zero, we considered the following confirmatory factor analysis model: $y_i = \mu + \Lambda \eta_i + \epsilon_i$, where y_i is the unobserved continuous counterpart of a set of observed ordinal indicators, U_i ; further, we assumed that $\eta_i \sim N(\mathbf{0}, \Psi_\zeta)$, and $\epsilon_i \sim N(\mathbf{0}, \Psi_e)$. The CFA with ordinal manifest indicators is a special case of Equation (2). For this illustration, eight categorical manifest variables were used to identify two latent variables for 300 subjects (i.e., $n = 300$). The continuous counterparts of the categorical variables, $y_{i1}, \dots, y_{i8}$, were transformed to ordered categorical observations $u_{i1}, \dots, u_{i8}$ with the thresholds values $\tau_j = (-3.0^*, -2.0, -1.0, 0.0, 0.5, 2.0^*)$ for $j = 1, \dots, 4$ and $(-1.0^*, -0.5, 0.0, 1.0, 1.5, 2.0^*)$ for $j = 5, \dots, 8$. The parameters marked with an asterisk were fixed at the values indicated for identification purposes. Other parameters were set to the values of $\mu = (0, \dots, 0)^T$, $\Psi_e = 0.8I_8$, and

$$\Lambda^T = \begin{bmatrix} 1.0^* & \lambda_{21} & \lambda_{31} & \lambda_{41} & 0.0^* & \underline{\lambda_{61}} & 0.0^* & 0.0^* \\ 0.0^* & \underline{\lambda_{22}} & 0.0^* & 0.0^* & 1.0^* & \lambda_{62} & \lambda_{72} & \lambda_{82} \end{bmatrix}, \quad (16)$$

in which $\lambda_{jk} = 0.8$ for $j = 2, 3, 4, 6, 7, 8$ and $k = 1, 2$. The true values of the cross-loadings, λ_{22} and λ_{61} (underlined in (16)) were set to 0.5. The factor covariance matrix was set to $\Psi_\zeta = \begin{pmatrix} 1.0 & -0.3 \\ -0.3 & 1.0 \end{pmatrix}$. Missing data were generated via the following logistic regression model: $\text{logit}[\text{pr}(r_{ij} = 1 | \boldsymbol{\phi})] = \phi_0 + \phi_1 y_{ij}$, with true parameter values $\phi_0 = -1.5$ and $\phi_1 = 0.4$. The proportion of missing data was about 0.185.

Based on: (1) the choices of conjugate priors assumed for μ , Λ_k , $\psi_{\epsilon k}^{-1}$ and Ψ_ζ ; (2)

simplification of the model due to the absence of changes in the latent factors over time (i.e., the vector of dynamic functions among latent factors, $f_t(\eta_{i,t-1}, \mathbf{b}_i, \boldsymbol{\theta}_\eta)$, is equal to $\mathbf{0}$, a vector of zeros); (3) the fact that $T_i = 1$ for all individuals; and (4) the lack of random effects in the model, the posterior distributions in the CFA illustration simplify considerably. For instance, steps (1) – (5) for sampling from the conditional distributions related to the non-parametric $\mathbf{b}_i$, as described in Supplementary Appendix A, are no longer needed. In addition, the posterior distribution of η_i is now, where $p(\eta_i | Y, \mu, \Lambda, \Psi_e) \sim N(\Omega_\eta \Lambda^T \Psi_e^{-1} (y_i - \mu), \Omega_\eta)$, where

$\Omega_\eta = (\Psi_\zeta^{-1} + \Lambda^T \Psi_\epsilon^{-1} \Lambda)^{-1}$. Other conditional distributions are as described in Supplementary Appendix A, except that all expressions involving computations over T_i now span only $T_i = 1$. All resultant conditional distributions from this model thus assume standard analytic forms, from which sampling can be obtained in relatively straightforward ways. In our supplementary materials, we provide our own Matlab code for performing estimation and computations of local influence measures. For didactic purposes, we also include code written in the JAGS (*Just Another Gibbs Sampler*; Plummer, 2003) program and complementary R code to compute local influence measures.

We fit the specified model with λ_{22} and λ_{61} incorrectly set to 0. We then computed local influence measures by imposing perturbations to the factor loading matrix as

$$\Lambda_\omega^T = \begin{bmatrix} 1.0 & \lambda_{21} & \lambda_{31} & \lambda_{41} & \omega_5 & \omega_6 & \omega_7 & \omega_8 \\ \omega_1 & \omega_2 & \omega_3 & \omega_4 & 1.0 & \lambda_{62} & \lambda_{72} & \lambda_{82} \end{bmatrix},$$

where $\omega = (\omega_1, \dots, \omega_8)^T$. Thus, $\omega^0 = (0, \dots, 0)^T$ represents no perturbation.

Because we are only considering perturbations to the factor loading matrix, the effects of the perturbations on the log perturbed joint density function, $\log p(D_{\text{com}}, \boldsymbol{\eta}|\omega)$, are contained only in the perturbed likelihood function given by

$$p(Y|\omega) = (2\pi)^{-np/2} |\Psi_\epsilon|^{-n/2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n (y_i - \mu - \Lambda_\omega \eta_i)^T \Psi_\epsilon^{-1} (y_i - \mu - \Lambda_\omega \eta_i) \right\}.$$

The first-order partial derivatives with respect to ω are

$$\begin{aligned} \frac{\partial \log p(Y|\omega)}{\partial \omega_j} &= \sum_{i=1}^n (y_{ij} - \mu_j - \Lambda_{\omega_j} \eta_i) \eta_{2i} / \psi_{\epsilon_j} \text{ for } j = 1, \dots, 4 \text{ and} \\ \frac{\partial \log p(Y|\omega)}{\partial \omega_j} &= \sum_{i=1}^n (y_{ij} - \mu_j - \Lambda_{\omega_j} \eta_i) \eta_{1i} / \psi_{\epsilon_j} \text{ for } j = 5, \dots, 8, \end{aligned} \quad (17)$$

where $\eta_i = (\eta_{1i}, \eta_{2i})^T$ and Λ_{ω_j} is the j th row of factor loading matrix Λ_ω . Differentiating the expressions in (17) with respect to ω again to yield $\partial_{\omega_j \omega_k}^2 \log p(D_{\text{com}}, \boldsymbol{\eta}|\omega)$ and subsequently taking the expectations of these second derivatives with respect to $p(Y, \boldsymbol{\eta}, \boldsymbol{\eta}|\omega)$, we obtain:

$$G(\omega^0) = \text{diag}(g_{11}, \dots, g_{44}, g_{55}, \dots, g_{88}), \quad (18)$$

where $g_{kk} = n\alpha_{0ek}\psi_{\zeta_0}^{22}/\{\beta_{0ek}(\rho_0 - q - 1)\}$ for $k = 1, \dots, 4$, and $g_{kk} = n\alpha_{0ek}\psi_{\zeta_0}^{11}/\{\beta_{0ek}(\rho_0 - q - 1)\}$ for $k = 5, \dots, 8$; $\psi_{\zeta_0}^{jj}$ is the j th ($j = 1, 2$) diagonal element of Ψ_{ζ_0} , a matrix of hyperparameters in the inverse-Wishart prior distribution for Ψ_ζ , α_{0ek} and

β_{0ek} are hyperparameters in the gamma distribution used as the prior distribution for Ψ_e^{-1} (for details see Supplementary Appendices A and B). Inspection of values of g_{kk} indicated that their values were simply functions of particular hyperparameter values, and essentially varied evenly across the 8 perturbations in ω (see Figure 1A). In addition, $G(\omega^0)$ already assumed a diagonal structure, suggesting that this initial perturbation scheme, without the need for any further transformations, already imposes largely comparable amounts of perturbation on all elements in the factor loading matrix.

Next, substituting values of $G(\omega^0)$ computed using (18) into the local influence Equations in (13) and (15), we computed first-order local influence measure based on the Bayes factor (FI_{BF}), second-order local influence measure based on ϕ -divergence, $SI_{D_{\phi, e_j}}$, and first-order

local influence measure using the posterior means of the square-root of the parameters in the measurement model (specifically, we used

$d(\theta) = [\sqrt{\mu_1}, \dots, \sqrt{\mu_8}, \sqrt{\psi_{\epsilon 1}}, \dots, \sqrt{\psi_{\epsilon 8}}, \sqrt{\lambda_{21}}, \sqrt{\lambda_{31}}, \sqrt{\lambda_{41}}, \sqrt{\lambda_{62}}, \sqrt{\lambda_{72}}, \sqrt{\lambda_{82}}]^T$), using summary statistics from the MCMC samples obtained during the model fitting process. These MCMC samples were obtained using 3000 additional iterations after 4000 burn-in iterations. Selected first and second-order local influence measures are displayed in Figures 1(B)–(D). The misspecifications in cross-loadings were correctly identified by means of the Bayes factor and posterior means local influence measures, but not the ϕ -divergence measure. Thus, even though the incorrect omission of the cross-loadings, λ_{22} and λ_{61} , led to substantial deviations in the posterior means of the measurement parameters, the overall deviations in joint posterior distribution in terms of ϕ -divergence are not sensitive enough to help identify the sources of the misspecification. The Bayes factor, in contrast, was sensitive enough to help identify the correct sources of model misspecification.

Illustration III: Detecting Misspecification in the Dynamic Form of a DFA

This illustration serves to demonstrate the utility of the local influence measures in helping to detect possible misspecification in the dynamic functions in the DFA model. To do so we simulated data using the previously published nonlinear model of emotions based on Equation (1), and with $b_{11i} \sim N(0.2, 0.05)$, $b_{12i} \sim N(0, 0.001)$, $b_{22i} \sim N(0.2, 0.05)$, and $b_{21i} \sim N(0, 0.001)$. Eight categorical manifest variables were used as indicators to identify two latent variables for 50 subjects over 25 time points for the first 10 subjects, 27 time points for the following 20 subjects, 30 time points for the following 10 subjects, and 35 time points for the last 10 subjects, i.e., $T_1 = \dots = T_{10} = 25$, $T_{11} = \dots = T_{30} = 27$, $T_{31} = \dots = T_{40} = 30$, and $T_{41} = \dots = T_{50} = 35$.

The continuous counterparts of the categorical variables, $y_{it1}, \dots, y_{it8}$, were transformed to ordered categorical observations $u_{it1}, \dots, u_{it8}$ with the thresholds values: $\tau_k = (-3.0^*, -2.0, -1.0, 0.0, 0.5, 2.0^*)$ for $k = 2, 3, 4$ and $(-1.0^*, -0.5, 0.0, 1.0, 1.5, 2.0^*)$ for $k = 5, \dots, 8$. Other parameters in the measurement model (see Equation (2)) were set to the values of $A = \mu = \mathbf{0}_8$, $c_{it} = 1$, $\Psi_e = 0.8I_8$, and

$$\Lambda^T = \begin{bmatrix} 1.0^* & \lambda_{21} & \lambda_{31} & \lambda_{41} & 0.0^* & 0.0^* & 0.0^* & 0.0^* \\ 0.0^* & 0.0^* & 0.0^* & 0.0^* & 1.0^* & \lambda_{62} & \lambda_{72} & \lambda_{82} \end{bmatrix}, \quad (19)$$

in which $\lambda_{ij} = 0.8$ for $i = 2, 3, 4, 6, 7, 8$ and $j = 1, 2$. The process noise covariance matrix was set to $\Psi_\zeta = \begin{pmatrix} 1.0 & -0.3 \\ -0.3 & 1.0 \end{pmatrix}$. Missing data were generated via the following logistic regression model: $\text{logit}\{\text{pr}(r_{ij} = 1|\varphi)\} = \varphi_0 + \varphi_1 y_{ij} + \varphi_2 y_{i,t-1,j}$, with the true parameter values $\varphi_0 = -1.5$, $\varphi_1 = 0.4$, and $\varphi_2 = 0.3$ for $j = 1, \dots, 8$, where $\boldsymbol{\varphi} = (\varphi_0, \varphi_1, \varphi_2)^T$. The proportion of missing data was about 0.203.

Then, we fit a DFA model assuming linear dynamic functions at the factor level to these data as: $\boldsymbol{\eta}_{it} = \begin{pmatrix} b_{11i} & b_{12i} \\ b_{21i} & b_{22i} \end{pmatrix} \boldsymbol{\eta}_{i,t-1} + \boldsymbol{\zeta}_{it}$, which is a special case of Equation (3), and a measurement model which is a special case of Equation (2). Thus, the fitted dynamic model was misspecified in postulating linear dynamics in the forms of $b_{12i}\boldsymbol{\eta}_{2i,t-1}$ and $b_{21i}\boldsymbol{\eta}_{1i,t-1}$.

To illustrate the utility of the proposed approach in detecting these sources of misspecification dynamic functions, we considered the following perturbation scheme:

$$\begin{aligned} \eta_{1it} &= \omega_1 b_{11i} \eta_{1i,t-1} + \omega_3 b_{12i} \eta_{2i,t-1} + \zeta_{1it} \\ \eta_{2it} &= \omega_4 b_{21i} \eta_{1i,t-1} + \omega_2 b_{22i} \eta_{2i,t-1} + \zeta_{2it} \end{aligned}$$

In this case, $\boldsymbol{\omega} = (\omega_1, \dots, \omega_4)^T$, and ω_3 and ω_4 corresponded to the perturbations linked to the misspecified components in the DFA model. Further, $\boldsymbol{\omega}^0 = (1, 1, 1, 1)^T$ represents no perturbation.

Let $\mathbf{B}_{it\omega} = \begin{pmatrix} \omega_1 b_{11i} & \omega_3 b_{12i} \\ \omega_4 b_{21i} & \omega_2 b_{22i} \end{pmatrix}$. Then, the perturbation scheme can be rewritten as

$$\boldsymbol{\eta}_{it} = \mathbf{B}_{it\omega} \boldsymbol{\eta}_{i,t-1} + \boldsymbol{\zeta}_{it}, \quad \boldsymbol{\zeta}_{it} \sim \mathcal{N}(\mathbf{0}, \Psi_\zeta)$$

Because we are only considering perturbations to the dynamic functions, the effects of the perturbations on the perturbed log joint density function, $\log p(D_{\text{com}}, \boldsymbol{\theta}|\boldsymbol{\omega})$, are contained only in:

$$\log p(\mathbf{H}|\mathbf{b}, \boldsymbol{\omega}) = -\frac{N}{2} \log |\Psi_\zeta| - \frac{1}{2} \sum_{i=1}^n \sum_{t=1}^{T_i} (\boldsymbol{\eta}_{it} - \mathbf{B}_{it\omega} \boldsymbol{\eta}_{i,t-1})^T \Psi_\zeta^{-1} (\boldsymbol{\eta}_{it} - \mathbf{B}_{it\omega} \boldsymbol{\eta}_{i,t-1}), \quad (20)$$

where $N = \sum_{i=1}^n T_i$.

To determine the appropriate amount of perturbations, we first computed

$$G(\omega^0) = (\rho_0 + 2r + 2) \sum_{i=1}^n \sum_{t=1}^{T_i} E \left\{ A_{i,t-1}^* \Psi_{\zeta 0}^{-1} A_{i,t-1}^* \right\}, \text{ where}$$

$$A_{i,t-1}^* = \begin{pmatrix} b_{11i} \eta_{1i,t-1} & 0 & b_{12i} \eta_{2i,t-1} & 0 \\ 0 & b_{22i} \eta_{2i,t-1} & 0 & b_{21i} \eta_{1i,t-1} \end{pmatrix},$$

and $E\{\cdot\}$ is the expectation taken with respect to the joint distribution of $\eta_{i,t-1}$ and $\mathbf{b}_i = (b_{11i}, b_{21i}, b_{12i}, b_{22i})^T$. Values of the diagonal elements of $G(\omega^0)$ are 685.19, 681.50, 0.016, 0.016, respectively. These values indicate that the perturbations to the random effects b_{12i} and b_{22i} impose large perturbations on the log of the perturbed joint density, $\log p(D_{\text{com}}, \boldsymbol{\theta}|\omega) = \log p(D_{\text{com}}, \boldsymbol{\theta}|\omega)$ compared to other components. Thus, unlike Illustration II, this illustration requires the use of an alternative perturbation scheme to impose a comparable amount of perturbations to all of the elements in $\mathbf{b}_i$.

We then defined a new perturbation scheme by defining $\tilde{\omega} = \omega^0 + G(\omega^0)^{1/2}(\omega - \omega^0)$ to yield $G(\tilde{\omega}) = I_4$. We computed local influence measures using the perturbation model with 2000 additional iterations after 3000 burn-in iterations. The same first and second-order local influence measures as those used in Illustration II, including local influence measures based on the Bayes factor, the ϕ -divergence and the posterior means of the parameter functions $d(\boldsymbol{\theta}) = (\sqrt{\mu_1}, \dots, \sqrt{\mu_8}, \sqrt{\psi_{\varepsilon_1}}, \dots, \sqrt{\psi_{\varepsilon_8}}, \sqrt{\lambda_{21}}, \sqrt{\lambda_{31}}, \sqrt{\lambda_{41}}, \sqrt{\lambda_{62}}, \sqrt{\lambda_{72}}, \sqrt{\lambda_{82}})^T$, are displayed in Table 1. Inspection of Table 1 indicates that all three local influence measures correctly identified the sources of misspecification in the DFA model, namely, in the cross-lag dependencies between the two latent factors.

Illustration IV: Linear DFA with Outliers in Categorical Responses, and Simultaneous Sensitivity Analysis of Multiple Components

To illustrate the utility of the proposed approach in detecting influential data cases while simultaneously enabling local influence analysis on multiple modeling elements, we generated data using the same linear DFA model as in Illustration III. However, as distinct from Illustration III, the true distributions for $\mathbf{b}_i = (b_{11i}, b_{21i}, b_{12i}, b_{22i})^T$ were specified to be: $b_{11i} \sim N(0.15, 0.001)$, $b_{21i} \sim N(0.15, 0.001)$, and $b_{12i} \sim N(-.15, .001)$ and $b_{22i} \sim N(-.15, .001)$ for $i = 1, \dots, 50$. In addition, missing data were generated via the following logistic regression model: $\text{logit}\{\text{pr}(r_{ijt} = 1|\boldsymbol{\phi})\} = \phi_0 + \phi_1 y_{ijt} + \phi_2 y_{i,t-1,j}$ with true parameters $\phi_0 = -1.5$, and $\phi_1 = 0.4$ and $\phi_2 = 0.3$ for $j = 1, \dots, 8$, where $\boldsymbol{\phi} = (\phi_0, \phi_1, \phi_2)^T$. Specifically, we generated a random number u from the uniform distribution, and the observation y_{ijt} was specified as missing if $u < \text{pr}(r_{ijt} = 1|\boldsymbol{\phi})$. The proportion of missing data was about 0.203.

To demonstrate the utility of the proposed approach in detecting outliers in categorical responses, we added two outliers in the simulated data set by changing u_{ijt} as $u_{ijt} + 1$ for $i = 4$ and 50 and $t = 1, \dots, T_i$ and $j = 1, \dots, 8$. In addition, we set u_{ijt} to 7 if the modified value of

u_{itj} was larger than 7. To detect the two outlying cases, we devised a simultaneous perturbation scheme to perturb individuals' manifest variables, the posited models, the DP distribution, and the prior distributions of parameters that may be affected by the level shifts caused by these outlying cases. Thus, we chose to simultaneously perturb: (i) the continuous counterparts of the categorical observed variables, y_{it} , with perturbation scheme: $y_{it\omega} = y_{it} + \omega_i \mathbf{1}_p$; (ii) the prior distribution of μ with $p(\mu_{\omega}) \sim N(\mu^0 + \omega_{\mu} \mathbf{1}_p, \Sigma_0)$; (iii) the prior distribution of Ψ_{ζ} with $p(\Psi_{\zeta\omega_{\zeta}}) \sim IW(\rho_0, \omega_{\zeta} \Psi_{\zeta_0})$; (iv) the prior distribution of φ with $p(\varphi_{\omega_{\varphi}}) \sim N(\varphi_{\varphi}^0, H_{\varphi}^0)$; (v) the prior distribution of μ_Z with $p(\mu_{Z\omega_{\tau}}) \sim N_{n_b}(\mu_{Z_0} + \omega_{\tau 2} \mathbf{1}_{n_b}, \Psi_{\mu_Z} / \omega_{\tau 1})$; (vi) the missing data coefficient, φ_1 ; (vii) the distribution of $\mathbf{Z}_g$ with perturbation scheme: $\mathbf{Z}_{g\omega_z} \stackrel{D}{=} N_{n_b}(\mu_Z + \omega_{z2} \mathbf{1}_{n_b}, \Psi_Z / \omega_{z1})$; and (viii) the prior distribution of α with $p(\alpha_{\omega_{\alpha}}) \sim \Gamma(a_1, \omega_{\alpha} a_2)$.

The effects of the perturbations on the log perturbed joint density function, $\log p(D_{\text{com}}, \boldsymbol{\theta} | \boldsymbol{\omega})$, are contained in:

$$\begin{aligned} \log p(Y, \mu, \Psi_{\zeta}, \mu_Z, r, Z, \alpha | \boldsymbol{\omega}) \propto & -\frac{1}{2} \sum_{i=1}^n \sum_{t=1}^{T_i} (y_{it} + \omega_i \mathbf{1}_p - \mu - \Lambda \eta_{it})^T \Psi_{\epsilon}^{-1} (y_{it} + \omega_i \mathbf{1}_p - \mu - \Lambda \eta_{it}) \\ & - \frac{1}{2} (\mu - \mu^0 - \omega_{\mu} \mathbf{1}_p)^T \Sigma_0^{-1} (\mu - \mu^0 - \omega_{\mu} \mathbf{1}_p) + \frac{\rho_0}{2} \log(|\omega_{\zeta} \Psi_{\zeta_0}|) \\ & - \frac{\omega_{\zeta}}{2} \text{tr}(\Psi_{\zeta_0} \Psi_{\zeta}^{-1}) - \frac{1}{2} (\varphi - \varphi_{\varphi}^0)^T (H_{\varphi}^0)^{-1} (\varphi - \varphi_{\varphi}^0) \\ & - \frac{1}{2} \log(|\omega_{\tau 1}^{-1} \Psi_{\mu_Z}|) - \frac{\omega_{\tau 1}}{2} (\mu_Z - \mu_{Z_0} - \omega_{\tau 2} \mathbf{1}_{n_b})^T \Psi_{\mu_Z}^{-1} (\mu_Z - \mu_{Z_0} - \omega_{\tau 2} \mathbf{1}_{n_b}) \\ & + \sum_{i=1}^n \sum_{t=1}^{T_i} \sum_{j=1}^p \left\{ r_{itj} (\varphi_0 + \omega_y \varphi_1 y_{itj} + \varphi_2 y_{i,t-1,j}) \right. \\ & \left. - \log(1 + \exp(\varphi_0 + \omega_y \varphi_1 y_{itj} + \varphi_2 y_{i,t-1,j})) \right\} \\ & - \frac{G}{2} \log|\Psi_Z / \omega_{z1}| - \frac{\omega_z}{2} \sum_{g=1}^G (\mathbf{Z}_g - \mu_Z - \omega_{z2} \mathbf{1}_{n_b})^T \Psi_Z^{-1} (\mathbf{Z}_g - \mu_Z - \omega_{z2} \mathbf{1}_{n_b}) \\ & + a_1 \log(\omega_{\alpha} a_2) - \omega_{\alpha} a_2 \alpha, \end{aligned}$$

where $\boldsymbol{\omega}_D = (\omega_1, \dots, \omega_n)^T$, $\boldsymbol{\omega}_P = (\omega_{\mu}, \omega_{\zeta}, \omega_{\varphi}, \omega_{\tau 1}, \omega_{\tau 2})^T$, and $\boldsymbol{\omega}_S = (\omega_y, \omega_{z1}, \omega_{z2}, \omega_a)^T$; $\boldsymbol{\omega} = (\boldsymbol{\omega}_D^T, \boldsymbol{\omega}_P^T, \boldsymbol{\omega}_S^T)^T$. Further, $\boldsymbol{\omega}_D^0 = \mathbf{0}_n$, $\boldsymbol{\omega}_P^0 = (0, 1, 1, 1, 0)^T$, and $\boldsymbol{\omega}_S^0 = (1, 1, 0, 1)^T$ represent no perturbation.

To determine the appropriate amount of perturbations, we first computed

$$G(\omega^0) = \text{diag}(g_{11}, \dots, g_{nn}, \mathbf{1}_p^T \Sigma_0^{-1} \mathbf{1}_p, \frac{\rho_0 q}{2}, (\varphi^0)^T (\mathbf{H}_\varphi^0)^{-1} \varphi^0, \frac{n_b}{2}, \mathbf{1}_{n_b}^T \Psi_{\mu_Z}^{-1} \mathbf{1}_{n_b}, g_{yy}, \frac{n_b G}{2}, \frac{G n_b c_1}{c_2}, a_1),$$

where $g_{ii} = T_i p \alpha_{0ek} / \beta_{0ek}$, and

$$g_{yy} = \sum_{i=1}^n \sum_{t=1}^{T_i} \sum_{j=1}^p E \left[\frac{\exp(\varphi_0 + \varphi_1 y_{itj} + \varphi_2 y_{i,t-1,j})}{(1 + \exp(\varphi_0 + \varphi_1 y_{itj} + \varphi_2 y_{i,t-1,j}))^2} \varphi_1^2 y_{itj}^2 \right], \text{ in which } E$$

denotes the expectation with respect to the priors of $\boldsymbol{\varphi}$ and y_{itj} . Note that α_{0ek} and β_{0ek} are hyperparameters in the gamma distribution used as the prior distribution for Ψ_ϵ^{-1} (see

Supplementary Appendices A and B), n_b is the dimension of random effects $\mathbf{b}_i$, q is the dimension of latent variables $\boldsymbol{\eta}_{ib}$, c_1 and c_2 are the hyperparameters in the gamma distribution used as the prior distribution for $\psi_{z_j}^{-1}$, a_2 is the hyperparameter in the gamma

distribution used as the prior distribution for α , and ρ_0 is the hyperparameter in the inverse-Wishart prior distribution for Ψ_ζ (see Supplementary Appendices A and B).

Values of the diagonal elements of $G(\omega^0)$ are shown in Figure 2(A). Inspection of values of g_{kk} indicated that individuals with more time points are associated with greater metric tensor values, namely, they are disproportionately perturbed compared to other cases. In addition, perturbations ω_y and ω_{z2} , namely, perturbation to the missing data coefficient, φ_1 and the mean of $\mathbf{Z}_g$, respectively, were disproportionate large compared to other perturbations. This highlights the importance of using an appropriate perturbation scheme.

As defined earlier, an “appropriate” perturbation scheme has the characteristics that all the $\mathcal{T}$ administered perturbations are orthogonal to each other, and of the same standardized magnitude. To accomplish this, we defined a new perturbation scheme by defining

$$\tilde{\omega} = \omega^0 + G(\omega^0)^{1/2} (\omega - \omega^0) \text{ to yield } \tilde{\omega}_i = \sqrt{g_{ii}} \omega_i \text{ for } i = 1, \dots, n,$$

$$\tilde{\omega}_{n+1} = \sqrt{\{\mathbf{1}_p^T \Sigma_0^{-1} \mathbf{1}_p\}} \omega_{n+1}, \tilde{\omega}_{n+5} = \sqrt{\{\mathbf{1}_{n_b}^T \Psi_{\mu_Z}^{-1} \mathbf{1}_{n_b}\}} \omega_{n+5}, \tilde{\omega}_{n+8} = \sqrt{\{Grc_1/c_2\}} \omega_{n+8},$$

$$\tilde{\omega}_{n+j} = 1 + \sqrt{g_j} (\omega_{n+j} - 1) \text{ for } j = 2, 3, 4, 6, 7, 9, \text{ where } g_2 = \rho_0 q / 2, g_3 = (\varphi^0)^T (\mathbf{H}_\varphi^0)^{-1} \varphi^0, g_4 = n_b / 2, g_6 = g_{yy}, g_7 = n_b G / 2 \text{ and } g_9 = a_1. \text{ This resulted in } \tilde{\omega} \text{ for which } G(\tilde{\omega}) = \mathbf{I}_{n+9}. \text{ That is,}$$

the different components of $\tilde{\omega}$ now imposed the same amount of perturbations on $\log p(\mathbf{Y}, \boldsymbol{\mu}, \Psi_\zeta, \boldsymbol{\mu}_Z, \mathbf{r}, \mathbf{Z}, \alpha | \boldsymbol{\omega})$ in various directions.

Selected first and second-order local influence measures are displayed in Figures 2(B)—(D), with a heuristic benchmark line obtained by adding 2 standard deviations of each local influence measure to its mean across the 59 perturbations considered. Inspection of Figures 2(B) and (C) indicates that subjects 4 and 50 were only correctly detected to be influential by the local influence measure constructed using posterior means of

$$d(\boldsymbol{\theta}) = (\sqrt{|\mu_1|}, \dots, \sqrt{|\mu_8|}, \sqrt{|\psi_{\epsilon 1}|}, \dots, \sqrt{|\psi_{\epsilon 8}|}, \sqrt{|\lambda_{21}|}, \sqrt{|\lambda_{31}|}, \sqrt{|\lambda_{41}|}, \sqrt{|\lambda_{62}|}, \sqrt{|\lambda_{72}|}, \sqrt{|\lambda_{82}|})^T; \text{ namely, by } \text{FI}_{M_d, \mathbf{e}_j} \text{ in}$$

plot (D). The local influence measures based on the ϕ -divergence and the Bayes factor both

missed subject 4 as an outlying case (see plots B and C). The local influence measure based on ϕ -divergence was close to identifying both outlying cases correctly in Figure 2(C), but was slightly underpowered under the heuristic benchmark line to detect subject 4 as an outlying case (note that the “outlying responses” were not allowed to exceed the 1-to-7 ordinal response scale, thus limiting the extremeness of these observed responses). In sum, this illustrative example demonstrates some of the inherent difficulties in detecting outlying cases in categorical responses, and the importance of considering multiple local influence measures in making diagnostic decisions about influential data points.

Illustration V: Detecting Misspecification in DP Prior and Missing Data Mechanism

This illustration serves to demonstrate the utility of the local influence measures in detecting the presence of multiple sources of model misspecification in the context of more complex dynamic models. To do so we generated data using the previously published nonlinear DFA model of emotions in Equation (1), and misspecified the base distribution of the nonparametric DP prior and the missing data mechanism. That is, we generated data using the same nonlinear DFA model as in Illustration III. However, as distinct from Illustration III, the true distributions for $\mathbf{b}_i = (b_{11i}, b_{21i}, b_{12i}, b_{22i})^T$ were specified to be: $b_{11i} \sim \text{Unif}(0.2, 0.4)$ for $i = 1, \dots, 25$, $b_{11i} \sim \text{Unif}(0.7, 0.8)$ for $i = 26, \dots, 50$, $b_{21i} \sim \text{Unif}(0.2, 0.4)$ for $i = 1, \dots, 25$, $b_{21i} \sim \text{Unif}(0.7, 0.8)$ for $i = 26, \dots, 50$, and $b_{12i} \sim N(-.15, .001)$ and $b_{22i} \sim N(-.15, .001)$ for $i = 1, \dots, 50$. In addition, missing data were generated via the following logistic regression model: $\text{logit}\{\text{pr}(r_{ijt} = 1 | \boldsymbol{\varphi})\} = \varphi_0 + \varphi_1 y_{ijt}$ with true parameter values $\varphi_0 = -2.0$, and $\varphi_1 = 1.2$ for $j = 1, \dots, 8$, where $\boldsymbol{\varphi} = (\varphi_0, \varphi_1)^T$. The proportion of missing data was about 0.16.

To summarize, two sources of misspecification were specified. First, we fit the model depicted in Equations (3) and (2) to the simulated data described above assuming that the missing data were MAR. That is, we set $\varphi_1 = 0$ to impose the MAR assumption and subsequently implemented a perturbation to the sampling distribution of the missing data coefficient, φ_1 . For illustrative purposes, φ_0 was fixed at the true value of 2.0. Second, the base distribution of the DP prior, P_0 , whose constituent values $\mathbf{Z}_{L_g}$ were distributed as $N_{n_b}[\boldsymbol{\mu}_Z, \boldsymbol{\Psi}_Z]$, was specified to be overly restrictive as $N_{n_b}[\boldsymbol{\mu}_Z, .5\boldsymbol{\Psi}_Z]$.

We chose a perturbation scheme to simultaneously perturb: (i) $\mathbf{Z}_{g\omega_z} \stackrel{D}{=} N_{n_b}[\boldsymbol{\mu}_Z, \boldsymbol{\Psi}_Z/\omega_z]$, (ii) the prior distribution of $\boldsymbol{\mu}_Z$, namely, $p(\boldsymbol{\mu}_{Z\omega_z}) \sim N(\omega_{\mu_{z2}} \boldsymbol{\mu}_{Z0}, \boldsymbol{\Psi}_{\mu_Z}/\omega_{\mu_{z1}})$, (iii) the missing data coefficient, φ_1 by specifying: $\text{logit}\{\text{pr}(r_{ijt}/\varphi_0, \omega_y)\} = \varphi_0 + \omega_y y_{ijt}$ for all subjects, time points and indicators, (iv) the prior distribution of $\boldsymbol{\varphi}$ with $p(\boldsymbol{\varphi}_{\omega_\varphi}) \sim N[\omega_\varphi \boldsymbol{\varphi}^0, \mathbf{H}_\varphi^0]$, and (v) the prior distribution of $\boldsymbol{\Psi}_\zeta$ with $p(\boldsymbol{\Psi}_{\zeta\omega_\zeta}) = \text{IW}(\rho_0, \omega_\zeta \boldsymbol{\Psi}_\zeta)$. The effects of the perturbations on the log perturbed joint density function, $\log p(D_{\text{com}}, \boldsymbol{\vartheta} | \boldsymbol{\omega})$, are contained in:

$$\begin{aligned}
& \log p(\Psi_{\zeta}, \mu_Z, r, \varphi | \omega) \\
& \propto -\frac{G}{2} \log |\Psi_Z / \omega_z| - \frac{\omega_z}{2} \sum_{g=1}^G (\mathbf{Z}_g - \mu_Z)^T \Psi_Z^{-1} (\mathbf{Z}_g - \mu_Z) \\
& - \frac{1}{2} \log (|\omega_{\mu_z 1}^{-1} \Psi_{\mu_Z}|) - \frac{\omega_{\mu_z 1}}{2} (\mu_Z - \omega_{\mu_z 2} \mu_{Z_0})^T \Psi_{\mu_Z}^{-1} (\mu_Z - \omega_{\mu_z 2} \mu_{Z_0}) \\
& + \sum_{i=1}^n \sum_{t=1}^{T_i} \sum_{j=1}^p [r_{itj}(\varphi_0 + \omega_y y_{itj}) - \log(1 + \exp(\varphi_0 + \omega_y y_{itj}))] \\
& - \frac{1}{2} (\varphi - \omega_{\varphi} \varphi^0) (\mathbf{H}_{\varphi}^0)^{-1} (\varphi - \omega_{\varphi} \varphi^0) + \frac{\rho}{2} \log (|\omega_{\zeta} \Psi_{\zeta_0}|) - \frac{\omega_{\zeta}}{2} \text{tr}(\Psi_{\zeta_0} \Psi_{\zeta}^{-1}),
\end{aligned} \tag{21}$$

where $\omega = (\omega_z, \omega_{\mu_z 1}, \omega_{\mu_z 2}, \omega_y, \omega_{\varphi}, \omega_{\zeta})^T$. Further, $\omega^0 = (1, 1, 1, 0, 1, 1)^T$ represents no perturbation, with $\omega_y^0 = 0$ because the fitted model (i.e., the model to be perturbed) assumed MAR. Our key interest was to evaluate whether the simultaneous perturbation scheme above was able to help us identify misspecification in the missing data mechanism with high specificity.

To determine the appropriate amount of perturbations, we first computed

$$G(\omega^0) = \text{diag}\left(\frac{Gr}{2}, \frac{n_b}{2}, \mu_{Z_0}^T \Psi_{\mu_Z}^{-1} \mu_{Z_0}, g_{yy}, \mu_{\varphi}^T \mathbf{H}_{0\varphi}^{-1} \mu_{\varphi}, \frac{\rho q}{2}\right),$$

where $g_{yy} = \sum_{i=1}^n \sum_{t=1}^T \sum_{j=1}^p E \left[\frac{\exp(\varphi_0 + y_{itj})}{(1 + \exp(\varphi_0 + y_{itj}))^2} y_{itj}^2 \right]$ and E denotes the expectation

with respect to the prior of φ and y_{itj} . Values of the diagonal elements of $G(\omega^0)$ are shown in Figure 3(A) to highlight the importance of devising an alternative perturbation scheme. Figure 3(A) indicates that the perturbations to the base distribution P_0 and the missing data coefficient φ_1 imposed disproportionately large perturbations on the log of the perturbed joint density compared to other components. We then defined a new perturbation scheme by defining $\tilde{\omega} = \omega^0 + G(\omega^0)^{1/2} (\omega - \omega^0)$ to yield $\tilde{\omega}_i = 1 + \sqrt{g_{ii}}(\omega_i - 1)$ and $G(\tilde{\omega}) = \mathbf{I}_6$. We computed local influence measures using the perturbation model with 3000 additional iterations after 5000 burn-in iterations.

Selected first and second-order local influence measures are displayed in Figures 3(B)—(D). Inspection of Figures 3(B) and (C) indicates that the local influence measures constructed using the Bayes factor and the ϕ -divergence correctly identified misspecifications in the distribution of $\mathbf{Z}_g$ and the MAR assumption. Figure 3(D) suggests that the local influence measure based on the posterior means of $d(\vartheta) = (\mu_1, \dots, \mu_8, \lambda_{21}, \lambda_{31}, \lambda_{41}, \lambda_{62}, \lambda_{72}, \lambda_{82}, \psi_{e1}, \dots, \psi_{e8})^T$ was not sensitive enough in this context as a measure for detecting

misspecification in the base DP distribution, although misspecification in the MAR assumption was correctly detected. Thus, supplementing Illustration IV, the present illustration reiterates the importance of considering diagnostic information from multiple local influence measures, and that the “best-performing” local influence measure may be model- and context-dependent.

Empirical Example Using Ecological Momentary Assessment (EMA) Data

As part of a larger study (see Diener et al., 1995), 174 college students (93 male and 81 female; average age = 20.24 years, SD = 1.41) were asked to provide daily ecological momentary assessment (EMA) ratings over 52 consecutive days. The participants were asked to indicate how often they had experienced a series of emotions since the last assessment using a 7-point Likert scale (with 1 = “None” to 7 = “Always”). Their responses on 8 ordinal items, namely, the items *joy*, *contentment*, *love*, *affection*, *unhappiness*, *anger*, *depression*, and *anxiety*, were used in the present application. The first four items were items that served to indicate an underlying unobserved (i.e., latent) variable, positive emotion (PE), while the last four items were as indicators of individuals’ latent negative emotion (denoted below as NE). The proportion of missing data in the sample was about 11.2%.

We fit a linear DFA model to the real data as in Illustration III. That is, we considered

$$f_t(\eta_{i,t-1}, \mathbf{b}_i, \boldsymbol{\theta}_\eta) = \begin{pmatrix} b_{11i} & b_{12i} \\ b_{21i} & b_{22i} \end{pmatrix} \eta_{i,t-1}. \text{ In the measurement model in (2), we considered } \mathbf{A} = (\mu_1, \dots, \mu_8)^T \text{ and } \mathbf{c}_{it} = 1.0, \text{ and a confirmatory structure of } \boldsymbol{\Lambda} \text{ with selected entries fixed at the values indicated in Equation (19), and no cross-loadings. To identify the parameters that were related to the ordered categorical variables in } \mathbf{y}_{it}, \text{ we set the first and last thresholds of all manifest ordinal items, } \tau_{jk} \text{ (} k = 1, 6; j = 1, \dots, 8), \text{ to } \Phi^{*-1}(g_{jk}), \text{ where } g_{jk} \text{ is the observed cumulative marginal proportion of responses for which } u_{ijt} < k, \text{ and } \Phi^{*-1} \text{ is the inverse standard normal cumulative distribution function. To accommodate possible presence of NMAR, we considered the following model}$$

$\dots, \mu_8)^T$ and $\mathbf{c}_{it} = 1.0$, and a confirmatory structure of $\boldsymbol{\Lambda}$ with selected entries fixed at the values indicated in Equation (19), and no cross-loadings. To identify the parameters that were related to the ordered categorical variables in $\mathbf{y}_{it}$, we set the first and last thresholds of all manifest ordinal items, τ_{jk} ($k = 1, 6; j = 1, \dots, 8$), to $\Phi^{*-1}(g_{jk})$, where g_{jk} is the observed cumulative marginal proportion of responses for which $u_{ijt} < k$, and Φ^{*-1} is the inverse standard normal cumulative distribution function. To accommodate possible presence of NMAR, we considered the following model

$$\text{logit}\left\{\text{pr}(r_{ijt} = 1 | \boldsymbol{\varphi})\right\} = \varphi_0 + \varphi_1 y_{ijt} + \varphi_2 y_{i,t-1,j} \text{ for } j = 1, \dots, 8.$$

That is, φ_1 and φ_2 were freely estimated (i.e. NMAR was assumed). Choices for all prior specifications, hyperparameters, and tuning parameters in the Gibbs sampler are summarized in Supplementary Appendix B.

We fixed the NMAR missing data coefficients associated with y_{ijt} and $y_{i,t-1,j}$ at zero, thereby imposing a MAR mechanism on the data. We then perturbed these coefficients with several other components whose inferential results might be affected by misspecification in the missing data mechanism. Four Bayesian local influence diagnostics were constructed using the Bayes factor, ϕ -divergence, and posterior mean local influence function computed using all parameters in the measurement model, and a posterior mean local influence function computed using all person-invariant parameters in the dynamic model, namely, those in the process noise covariance matrix, $\boldsymbol{\Psi}_\zeta$. That is, for the first posterior mean function, we took $\mathbf{d}(\boldsymbol{\theta}) = (\sqrt{|\mu_1|}, \dots, \sqrt{|\mu_8|}, \sqrt{|\psi_{\varepsilon 1}|}, \dots, \sqrt{|\psi_{\varepsilon 8}|}, \sqrt{|\lambda_{21}|}, \sqrt{|\lambda_{31}|}, \sqrt{|\lambda_{41}|}, \sqrt{|\lambda_{62}|}, \sqrt{|\lambda_{72}|}, \sqrt{|\lambda_{82}|})^T$. For the second

posterior mean function we took $\mathbf{d}(\boldsymbol{\theta}) = (\Psi_{\zeta 11}, \Psi_{\zeta 12}, \Psi_{\zeta 22})^T$, where $\Psi_{\zeta jk}$ is the (j, k) th component of matrix Ψ_{ζ} for $j, k = 1, 2$. Three separate chains with distinct starting values were used to compute EPSR values for determining the stability of parameter estimates across chains; a total of 10000 burn-in iterations were used, with 5000 additional iterations used for obtaining samples for other computational and inferential purposes after burn-in.

The perturbed joint posterior density function is given by

$$\log p(Y, \boldsymbol{\mu}, \Psi_{\zeta}, \boldsymbol{\mu}_Z, \mathbf{r}, \mathbf{H}, \boldsymbol{\alpha} | \boldsymbol{\omega}) \quad (22)$$

$$\begin{aligned} & \propto -\frac{1}{2} \sum_{i=1}^n \sum_{t=1}^{T_i} (\mathbf{y}_{it} + \omega_i \mathbf{1}_p - \boldsymbol{\mu} - \Lambda \boldsymbol{\eta}_{it})^T \Psi_{\varepsilon}^{-1} (\mathbf{y}_{it} + \omega_i \mathbf{1}_p - \boldsymbol{\mu} - \Lambda \boldsymbol{\eta}_{it}) \\ & - \frac{1}{2} (\boldsymbol{\mu} - \boldsymbol{\mu}^0 - \omega_{\mu} \mathbf{1}_p)^T \Sigma_0^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}^0 - \omega_{\mu} \mathbf{1}_p) + \frac{\rho_0}{2} \log(l \omega_{\zeta} \Psi_{\zeta_0} |) \\ & - \frac{\omega_{\zeta}}{2} \text{tr}(\Psi_{\zeta_0} \Psi_{\zeta}^{-1}) - \frac{1}{2} (\boldsymbol{\varphi} - \omega_{\varphi} \boldsymbol{\varphi}^0)^T (\mathbf{H}_{\varphi}^0)^{-1} (\boldsymbol{\varphi} - \omega_{\varphi} \boldsymbol{\varphi}^0) \\ & - \frac{1}{2} \log(l \omega_{\tau 1}^{-1} \Psi_{\mu_Z} |) - \frac{\omega_{\tau 1}}{2} (\boldsymbol{\mu}_Z - \boldsymbol{\mu}_{Z_0} - \omega_{\tau 2} \mathbf{1}_{n_b})^T \Psi_{\mu_Z}^{-1} (\boldsymbol{\mu}_Z - \boldsymbol{\mu}_{Z_0} - \omega_{\tau 2} \mathbf{1}_{n_b}) \\ & + \sum_{i=1}^n \sum_{t=1}^{T_i} \sum_{j=1}^p \left\{ r_{ij} (\varphi_0 + \omega_{y1} \varphi_1 y_{ij} + \omega_{y2} \varphi_2 y_{i,t-1,j}) \right. \\ & \left. - \log(1 + \exp(\varphi_0 + \omega_{y1} \varphi_1 y_{ij} + \omega_{y2} \varphi_2 y_{i,t-1,j})) \right\} \\ & - \frac{G}{2} \log |\Psi_Z / \omega_{z1}| - \frac{\omega_{z1}}{2} \sum_{g=1}^G (\mathbf{Z}_g - \boldsymbol{\mu}_Z - \omega_{z2} \mathbf{1}_{n_b})^T \Psi_Z^{-1} (\mathbf{Z}_g - \boldsymbol{\mu}_Z - \omega_{z2} \mathbf{1}_{n_b}) \\ & + a_1 \log(\omega_{\alpha} a_2) - \omega_{\alpha} a_2 \alpha - \frac{1}{2} \sum_{i=1}^n \sum_{t=1}^{T_i} (\boldsymbol{\eta}_{it} - \mathbf{B}_{it\omega} \boldsymbol{\eta}_{i,t-1})^T \Psi_{\zeta}^{-1} (\boldsymbol{\eta}_{it} - \mathbf{B}_{it\omega} \boldsymbol{\eta}_{i,t-1}) \end{aligned}$$

$$\text{where } \mathbf{B}_{it\omega} = \begin{pmatrix} \omega_{b1} b_{11i} & \omega_{b3} b_{12i} \\ \omega_{b4} b_{21i} & \omega_{b2} b_{22i} \end{pmatrix}, \boldsymbol{\omega}_D = (\omega_1, \dots, \omega_n)^T, \boldsymbol{\omega}_P = (\omega_{\mu}, \omega_{\zeta}, \omega_{\varphi}, \omega_{\tau 1}, \omega_{\tau 2})^T, \boldsymbol{\omega}_S =$$

$(\omega_{y1}, \omega_{y2}, \omega_{z1}, \omega_{z2}, \omega_{\alpha}, \omega_{b1}, \omega_{b2}, \omega_{b3}, \omega_{b4})^T$, and $\boldsymbol{\omega} = (\omega_D^T, \omega_P^T, \omega_S^T)^T$. Furthermore, $\boldsymbol{\omega}_D^0 = \mathbf{0}_n$, $\boldsymbol{\omega}_P^0 = (0, 1, 1, 1, 0)^T$, and $\boldsymbol{\omega}_S^0 = (1, 1, 1, 0, 1, 1, 1, 1, 1)^T$ represent no perturbations. To determine the appropriate amount of perturbations, we first computed

$$\mathbf{G}(\boldsymbol{\omega}^0) = \begin{pmatrix} \mathbf{G}_D(\boldsymbol{\omega}^0) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_P(\boldsymbol{\omega}^0) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_S(\boldsymbol{\omega}^0) \end{pmatrix},$$

in which $\mathbf{G}_D(\boldsymbol{\omega}^0) = \text{diag}(g_{11}, \dots, g_{nn})$, where $g_{ii} = T_i \rho a_{0ek} / \beta_{0ek}$;

$\mathbf{G}_P(\boldsymbol{\omega}^0) = \text{diag}\left(\mathbf{1}_p^T \boldsymbol{\Sigma}_0^{-1} \mathbf{1}_p, \frac{\rho_0 q}{2}, (\boldsymbol{\varphi}^0)^T (\mathbf{H}_{\boldsymbol{\varphi}}^0)^{-1} \boldsymbol{\varphi}^0, \frac{n_b}{2}, \mathbf{1}_{n_b}^T \boldsymbol{\Psi}_{\mu_Z}^{-1} \mathbf{1}_{n_b}, g_{yy1}, g_{yy2}, \frac{rG}{2}, \frac{Grc_1}{c_2}, a_1\right)$, where

$$g_{yy1} = \sum_{i=1}^n \sum_{t=1}^{T_i} \sum_{j=1}^p E \left[\frac{\exp(\varphi_0 + \varphi_1 y_{itj} + \varphi_2 y_{i,t-1,j}) \varphi_1^2 y_{itj}^2}{(1 + \exp(\varphi_0 + \varphi_1 y_{itj} + \varphi_2 y_{i,t-1,j}))^2} \right] \text{ and}$$

$$g_{yy2} = \sum_{i=1}^n \sum_{t=1}^{T_i} \sum_{j=1}^p E \left[\frac{\exp(\varphi_0 + \varphi_1 y_{itj} + \varphi_2 y_{i,t-1,j}) \varphi_2^2 y_{i,t-1,j}^2}{(1 + \exp(\varphi_0 + \varphi_1 y_{itj} + \varphi_2 y_{i,t-1,j}))^2} \right]; \text{ and } E$$

denotes the expectation with respect to the priors of $\boldsymbol{\varphi}$ and the distributions of y_{itj} 's. In

addition, $\mathbf{G}_S(\boldsymbol{\omega}^0) = (\rho_0 + 2r + 2) \sum_{i=1}^n \sum_{t=1}^{T_i} E_A \left[\mathbf{A}_{i,t-1}^* \boldsymbol{\Psi}_{\zeta_0}^{-1} \mathbf{A}_{i,t-1}^* \right]$, where

$$\mathbf{A}_{i,t-1}^* = \begin{pmatrix} b_{11i} \eta_{1i,t-1} & 0 & b_{12i} \eta_{2i,t-1} & 0 \\ 0 & b_{22i} \eta_{2i,t-1} & 0 & b_{21i} \eta_{1i,t-1} \end{pmatrix},$$

and $E_A\{\cdot\}$ is the expectation taken with respect to the joint distribution of $\boldsymbol{\eta}_{i,t-1}$ and $\mathbf{b}_i = (b_{11i}, b_{22i}, b_{12i}, b_{21i})^T$.

The metric tensor values $g_{ik}(\boldsymbol{\omega}^0)$ and the three local influence measures are plotted in Figure 4(A)–(D). Due to the equal number of measurement occasions across individuals, the metric tensor values $g_{ik}(\boldsymbol{\omega}^0)$ are constant for all the individuals $i = 1, \dots, n$ in the study, as shown in Figure 4(A). In addition, large deviations in the metric tensor value (indicating a large perturbation) was associated with $\omega_{b1} - \omega_{b3}$, namely, perturbations to the person-specific auto- and cross-regression parameters.

Based on the three Bayesian local influence measures and the corresponding benchmark lines (obtained as the mean of a particular measure across individual case perturbations ± 2 standard deviations of the measure), cases 8, 105, 128, 149, and 156 were detected to be influential by means of the Bayes factor local influence measure, $\text{FI}_{\text{BF}, e_j}$ (see Figure 4(B)),

as well as the posterior mean local influence measures constructed using the measurement parameters (see Figure 4(D)). The local influence measure constructed using the posterior mean function of all dynamic parameters yielded similar results and the corresponding plot is omitted to conserve space. All of the same cases except for case 156 were also detected to be influential by means of the ϕ -divergence measure, $\text{SI}_{\text{D}, \phi, e_j}$ (see Figure 4(C)). All local

influence measures identified ω_{y1} and ω_{y2} , the perturbations linked to the two NMAR coefficients whose values were incorrectly set to the value of 0, to be influential.

We fit a second DFA model after removing the outlying cases noted above, and with the NMAR missing data coefficients linked to y_{it} and $y_{i,t-1}$, φ_1 and φ_2 , freely estimated. Inspection of the local influence measures generated from the revised analysis suggested that only ω_{y1} remained to be influential. The estimates from this revised model are shown in Table 2. The credible intervals of the NMAR coefficients, φ_1 and φ_2 , did not include 0, and

were positive (posterior mean on $\phi_1 = 2.331$, 95% credible interval = [2.152, 2.481]) and negative (posterior mean on $\phi_2 = -0.409$, 95% credible interval = [-0.513, -0.301]) in sign, respectively. This suggested that the data were not well characterized by an MAR model, and that higher underlying affect ratings on the same day tended to increase the logodds of missing data on that day, whereas higher underlying affect ratings from the previous day tended to lower the logodds of missingness today.

Parameters whose standardized differences in posterior means before and after the outlier removal exceeded 2 in absolute value are underlined and highlighted in bold in Table 2, with positive (negative) standardized differences reflecting increases (decreases) in posterior mean estimates following the removal of outlying cases and the MAR assumption. Most of the parameters that showed notable changes were parameters associated with NE and its indicators. The estimate for process noise variance for NE, for instance, showed clear increase in magnitude after the removal of the outlying cases and MAR assumption, suggesting that the revised model and screened data were able to capture more variability in NE at the latent level than the original model and data. For the individual-specific time series parameters, we computed the 95% highest posterior density (HPD) intervals of their distributions using the empirical shortest probability interval approach (Chen and Shao, 1999), and summarized them in Table 2. At the group level, substantive conclusions concerning the time series parameters remained largely unaltered from the original model to the revised model, with the exception that the 95% HPDs for the posterior means of b_{21j} , the PE $\rightarrow$ NE cross-regression effect, went from including zero (95% HPD interval = [-0.02, 0.96]) to now excluding zero (95% HPD interval = [0.11, 1.06]).

Overall, our results highlighted the sensitivity of the measurement parameters to outlying cases and assumptions concerning the missing data mechanism. The person-specific time series parameters for the latent factors were also sensitive to perturbations to the missing data assumptions, with b_{21j} , the PE $\rightarrow$ NE cross-regression effect appeared unstable and sensitive to changes in the missing data assumption and outlying cases. In addition, perturbation to ϕ_1 in the revised NMAR model was found to be influential on all the local influence measures considered, thus suggesting the need for a formal investigation of the missing data mechanism underlying both emotions, and the tenability of constraining the magnitude of ϕ_1 and ϕ_2 to be invariant across PE and NE in predicting individuals' logodds of missingness. Overall, our results highlight the importance of performing sensitivity analysis on dynamic modeling results involving individuals' emotions, especially when the data may be NMAR. Given the sensitivity of the modeling results to variations in missing data assumptions in the present example, interpretations have to be performed with caution, and more thorough consideration of ways to better handle the missing data to improve the robustness of the modeling results is warranted.

Discussion

We utilized a Bayesian local influence method to conduct sensitivity analysis of DFA models with categorical data, nonignorable missing outcomes and nonparametric random effect distributions for the time series parameters at the latent level. Four illustrations and

one empirical application were included to demonstrate the utility of the proposed approach as a sensitivity and outlier/model misspecification detection tool.

Results from the empirical local influence analysis suggested that individuals' PE and NE showed evidence of NMAR. Modeling results pertaining to the cross-regression effects between PE and NE appeared to be sensitive to variations in the missing data model, thus underscoring the importance of performing sensitivity analysis in EMA studies of emotions. Other future extensions may include construction of perturbation schemes to assess sensitivity of modeling results to missingness in the covariates, as well as changes in the specification of the NMAR mechanisms or functional forms of the dynamic functions for PE and NE. Furthermore, in the empirical as well as other illustrations, we used empirical means and standard deviations of the local influence measures to construct benchmark lines for identifying especially influential perturbations. Such benchmark lines are, of course, arbitrary and are dependent on the perturbations pursued. Future studies should utilize approaches such as bootstrapping (Zhu et al., 2012) to construct empirical distribution for the local influence measures to identify components whose influence exceeds some critical thresholds determined using the bootstrap samples.

In the approach utilized here, we used local influence analysis in the context of various special cases of DFA models to identify outlying cases and possible sources of model misspecification. In practice, to better understand the specific ways in which outlying cases may show deviations in dynamics compared to the sample as a whole, other case perturbations may have to be considered. Perturbing each individual's random effects vector separately, and identifying individual cases for whom the assumption of MAR may not be tenable through individual-specific perturbations of the NMAR coefficients, ϕ_1 and ϕ_2 , are among some of the possibilities for probing individual differences.

The DP prior was used in the present article as a nonparametric prior for the distribution of random effects. Taking a nonparametric or semiparametric approach allows the assumption of normality to be treated as a testable hypothesis, but also brings up new questions concerning the adequacies and robustness of such nonparametric priors – of which the DP prior adopted here is but one special case – in tailoring to various distributional characteristics. Making such decisions is particularly challenging because “conventional” modeling routes based on comparing models that assume different prior densities are not feasible in this case. That is, it is computationally formidable to obtain model fit indices (e.g., Bayes factor) comparing models with different prior and sampling distribution assumptions. Sensitivity analysis appears to provide a viable, alternative way to assess the adequacies of the DP prior at least locally within the context of the current model formulations. In the context of the empirical application, we found that our modeling results remained relatively robust to local perturbations to the DP prior, but without a flexible local influence approach such as that adopted in the present study, simultaneous evaluation of the DP-related assumptions and other perturbations would have been computationally prohibitive.

We only considered three objective functions in constructing the local influence measures in the present study. Such choices are by no means exhaustive. Other objective functions such

as posterior standard deviations, HPD intervals, and other distance measures are all viable alternatives. Our illustrations demonstrated that the optimal and “best-performing” local influence measure may vary across modeling contexts and scenarios, thereby highlighting the importance of considering multiple local influence measures in making diagnostic decisions. Each of the local influence measures considered has its strengths, but also some limitations. For instance, the Bayes factor-based local influence measure may require reasonably informative priors to demonstrate sufficient sensitivity in detecting model misspecification (Berger, 2004). The ϕ -divergence measure is known to fail in cases where the reference (i.e., the unperturbed joint posterior) density is equal to zero, thus requiring either some adaptations to the fitted model or corrections to the divergence function for the ϕ -divergence measure to be used in such cases. In a similar vein, the posterior mean-based local influence measure may be informative in some contexts, but may not provide effective quantification of changes in modeling uncertainties under different perturbations.

In closing, we note that even though Bayesian analytic methods have received increasing attention within the field of psychometrics in the last few decades, its popularity in nowhere parallels the proliferation of Bayesian data analytic techniques in the statistical literature. This is due in part to some psychometricians’ reservations over the seemingly “unlimited” amount of flexibility afforded by Bayesian and the related Markov chain Monte Carlo methods, and the lack of rigorous methods for performing model diagnostics and assumption checking in more complex models. The proposed local influence approach is among the repertoire of tools that can help researchers investigate, and possibly alleviate the sources of such reservations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Frank Fujita for permission to use his published data set. Funding for this study was provided by NSF grant SES-1357666, NIH grant R01GM105004 and NSFC grant 11671349.

Appendix

Local Influence Measures Based on Bayes Factor, Posterior Means, and ϕ -Divergence **Bayes factor.** The Bayes factor can be regarded as a statistic for comparing the change in model fit in the presence of ω relative to ω^0 . Using the Bayes factor as the objective function and setting $\mathbf{W}_f = \mathbf{I}_T$, the key components needed to compute first-order local influence measures using the Bayes factor are $G(\omega^0)$ and $\nabla_{BF} = \omega BF(\omega^0)$, namely the changes in the Bayes factor with respect to the perturbations imposed by ω near ω^0 . The latter can be computed as: (Zhu et al., 2011, 2014):

$$\begin{aligned}
\nabla_{BF} &= \partial_{\omega} \left\{ \log p(D_{\text{obs}} | \omega) - \log p(D_{\text{obs}} | \omega^0) \right\}_{\omega = \omega^0} \quad (23) \\
&= \partial_{\omega} \left\{ \log \left(\int p(D_{\text{com}} | \boldsymbol{\vartheta}, \omega) p(\boldsymbol{\vartheta} | \omega) dD_{\text{mis}} d\boldsymbol{\vartheta} \right) \right\}_{\omega = \omega^0} \\
&= E_{\omega^0} [\partial_{\omega} \log p(D_{\text{com}}, \boldsymbol{\vartheta} | \omega^0) D_{\text{obs}}, \omega^0]
\end{aligned}$$

where E_{ω^0} denotes expectation taken with respect to $p(D_{\text{mis}}, \boldsymbol{\vartheta} | D_{\text{obs}}, \omega^0)$. This expectation may be approximated by averaging over the MCMC samples, $\{(D_{\text{mis}}^{(\kappa)}, \boldsymbol{\vartheta}^{(\kappa)}) : \kappa = K_0 + 1, \dots, K_1\}$, as:

$$\nabla_{BF} \approx K_1^{-1} \sum_{\kappa = K_0 + 1}^{K_0 + K_1} \partial_{\omega} \log p(D_{\text{obs}}, D_{\text{mis}}^{(\kappa)}, \boldsymbol{\vartheta}^{(\kappa)} | \omega^0). \quad (24)$$

Posterior mean

Let $d(\boldsymbol{\vartheta})$ denote a vector of functions of $\boldsymbol{\vartheta}$ of interest from an inferential standpoint (consisting of e.g., square-root of variance parameters) and setting $W_f \equiv W_{M_d} = [\text{Cov}(d(\boldsymbol{\vartheta}) | D_{\text{obs}})]^{-1}$, the inverse of the posterior covariance matrix of $d(\boldsymbol{\vartheta})$. One may construct another first-order local influence measure using the posterior means of $d(\boldsymbol{\vartheta})$ as another objective function. As summarized elsewhere (Zhu et al., 2014) ∇_{M_d} can be computed as:

$$\begin{aligned}
\nabla_{M_d} &= \text{Cov} \left\{ \partial_{\omega} \log p(D_{\text{com}}, \boldsymbol{\vartheta} | \omega^0), d(\boldsymbol{\vartheta}) | D_{\text{obs}} \right\} \quad (25) \\
&\approx K_1^{-1} \sum_{\kappa = K_0 + 1}^{K_0 + K_1} \left\{ d(\boldsymbol{\vartheta}^{(\kappa)}) \partial_{\omega} \log p(D_{\text{obs}}, D_{\text{mis}}^{(\kappa)}, \boldsymbol{\vartheta}^{(\kappa)} | \omega^0) \right\} \\
&\quad - [K_1^{-1} \sum_{\kappa = K_0 + 1}^{K_0 + K_1} d(\boldsymbol{\vartheta}^{(\kappa)})] [K_1^{-1} \sum_{\kappa = K_0 + 1}^{K_0 + K_1} \partial_{\omega} \log p(D_{\text{obs}}, D_{\text{mis}}^{(\kappa)}, \boldsymbol{\vartheta}^{(\kappa)} | \omega^0)],
\end{aligned} \quad (26)$$

where the elements marked with (κ) are MCMC samples used to approximate the covariance function in Equation (25).

ϕ -divergence

The ϕ -divergence between the joint posterior distributions with and without perturbations (i.e., $D_\phi(\omega) = E_{\omega^0} \left\{ \phi(p(D_{\text{mis}}, \boldsymbol{\theta} | D_{\text{obs}}, \omega) / p(D_{\text{mis}}, \boldsymbol{\theta} | D_{\text{obs}})) \right\}$) can be used as another objective function. It can be shown that $\nabla_{D_\phi} = \mathbf{0}$. Thus, to quantify local influence we resort to the second-order local influence measure and its approximation using simulated samples from the Gibbs sampler. Specifically, we compute the second-order changes in ϕ -divergence with respect to ω around ω^0 , $H_{D_\phi} = \partial_\omega^2 D_\phi(\omega^0)$ as:

$$\begin{aligned} H_{D_\phi} &= \ddot{\phi}(1) E_{\omega^0} \{ [\partial_\omega \log p(D_{\text{mis}}, \boldsymbol{\theta} | D_{\text{obs}}, \omega^0)]^{\otimes 2} | D_{\text{obs}}, \omega^0 \}, \\ &\approx \ddot{\phi}(1) \\ &\quad \frac{1}{K_1} \sum_{\kappa=K_0+1}^{K_0+K_1} \left\{ \partial_\omega \log p(D_{\text{obs}}, D_{\text{mis}}^{(\kappa)}, \boldsymbol{\theta}^{(\kappa)} | \omega^0) - \frac{1}{K_1} \sum_{\kappa'=K_0+1}^{K_0+K_1} \partial_\omega \log p(D_{\text{obs}}, D_{\text{mis}}^{(\kappa')}, \boldsymbol{\theta}^{(\kappa')} | \omega^0) \right\}^{\otimes 2}, \end{aligned} \quad (27)$$

where $q^{\otimes 2} = qq^T$ for any vector $q \in \mathbb{R}^{\mathcal{T}}$, $\ddot{\phi}(a) = \partial^2 \phi / \partial a^2$, and $\phi(\cdot)$ is a convex function, selected in the present study to be $\phi(a) = -\log(a)$ to reflect the Kullback-Leibler divergence (Kass et al., 1989; Weiss, 1996) between the posterior distributions with and without perturbation.

References

- Amari, SI. Differential-Geometrical Methods in Statistics Lecture Notes in Statistics 28. 2nd. Springer-Verlag; Berlin: 1990.
- Ansari A, Iyengar R. Semiparametric Thurstonian models for recurrent choices: A Bayesian analysis. *Psychometrika*. 2006; 71(4):631–657.
- Antoniak CE. Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *The Annals of Statistics*. 1974; 2(6):1152–1174.
- Berger JO. An overview of robust bayesian analysis. *Test*. 1994; 3:5–58.
- Berger, JO. Encyclopedia of Statistical Sciences. John Wiley & Sons, Inc; 2004. Bayes Factors; p. 00101
- Blozis SA, Ge X, Xu S, Natsuaki MN, Shaw DS, Neiderhiser JM, Scaramella LV, Leve LD, Reiss D. Sensitivity analysis of multiple informant models when data are not missing at random. *Structural Equation Modeling: A Multidisciplinary Journal*. 2013; 20(2):283–298. [PubMed: 25221420]
- Browne, MW., Nesselroade, JR. Representing psychological processes with dynamic factor models: Some promising uses and extensions of autoregressive moving average time series models. In: Maydeu-Olivares, A., McArdle, JJ., editors. *Contemporary psychometrics: A Festschrift for Roderick P. McDonald*. Erlbaum; Mahwah, NJ: 2005. p. 415–452.

- Chen MH, Shao QM. Monte Carlo estimation of Bayesian credible and HPD intervals. *Journal of Computational and Graphical Statistics*. 1999; 8:69–92.
- Chow SM, Nesselroade JR, Shifren K, McArdle JJ. Dynamic structure of emotions among individuals with Parkinson's disease. *Structural Equation Modeling*. 2004; 11:560–582.
- Chow SM, Tang N, Ibrahim JG, Zhu H. Nonlinear dynamic factor analysis model with nonignorable missingness for ecological momentary assessment affect data. 2017 under review.
- Chow SM, Tang N, Yuan Y, Song X, Zhu H. Bayesian estimation of semiparametric dynamic latent variable models using the dirichlet process prior. *British Journal of Mathematical and Statistical Psychology*. 2011; 64(1):69–106. [PubMed: 21506946]
- Copas J, Eguchi S. Local model uncertainty and incomplete data bias (with discussion). *Journal of the Royal Statistical Society Series B*. 2005; 67:459–512.
- Daniels, MJ., Hogan, JW. *Missing Data in Longitudinal Studies: Strategies for Bayesian Modeling and Sensitivity Analysis*. Chapman and Hall; London: 2008.
- De Beuckelaer, A., Swinnen, G. Biased latent variable mean comparisons due to measurement non-invariance: a simulation study. In: Davidov, E., Schmidt, P., Billiet, J., editors. *Cross-cultural analysis : methods and applications*. Taylor & Francis Group; 2011. p. 117–148. *European association of methodology series*
- Dey, DK., Ghosh, SK., Lou, KR. On local sensitivity measures in bayesian (with discussion). Berger, JO., Betto, B., Moreno, E., Pericchi, LR., Ruggeri, F., Salinetti, G., Wasserman, L., editors. Vol. 29. *Institute of Mathematical Statistics; Hayward, CA: 1996. p. 21–39. IMS Lecture Notes-Monograph Series*
- Diener E, Fujita F, Smith H. The personality structure of affect. *Journal of Personality and Social Psychology*. 1995; 69(1):130–141.
- Duncan KA, MacEachern SN. Nonparametric Bayesian modelling for item response. *Statistical Modeling*. 2008; 8(1):41–66.
- Dunson, DB. *Nonparametric Bayes Statistical Modeling*. Cambridge University Press; New York: 2008. *Nonparametric Bayes applications to biostatistics*.
- Engle RF, Watson M. A one-factor multivariate time series model of metropolitan wage rates. *Journal of the American Statistical Association*. 1981; 76:774–781.
- Ferguson Thomas S. A Bayesian analysis of some nonparametric problems. *The Annals of Statistics*. 1973; 1(2):209–230.
- Ferrer E, Nesselroade JR. Modeling affective processes in dyadic relations via dynamic factor analysis. *Emotion*. 2003; 3:344–360. [PubMed: 14674828]
- Forni M, Hallin M, Lippi M, Reichlin L. The generalized dynamic factor model: One-sided estimation and forecasting. *Journal of the American Statistical Association*. 2005; 100:830–840.
- Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Statist Sci*. 1992; 7(4):457–472.
- Geweke, JF. The dynamic factor analysis of economic time-series models. In: Aigner, D., Goldberger, A., editors. *Latent Variables in Soci-Economic Models*. North-Holland; New York: 1977. p. 365–383.
- Geweke JF, Singleton KJ. Maximum likelihood confirmatory factor analysis of economic time series. *International Economic Review*. 1981; 22:133–137.
- Gustafson P. Local sensitivity of inferences to prior marginals. *Journal of the American Statistical Association*. 1996a; 91:774–781.
- Gustafson P. Local sensitivity of posterior expectations. *Annals of Statistics*. 1996b; 24:174–195.
- Hens N, Aerts M, Molenberghs G, Thijs H, Verbeke G. Kernel weighted influence measures. *Computational Statistics and Data Analysis*. 2006; 48:467–487.
- Ibrahim JG, Chen MH, Lipsitz SR. Monte carlo EM for missing covariates in parametric regression models. *Biometrics*. 1999; 55:591–596. [PubMed: 11318219]
- Ishwaran, H., James. *Journal of the American Statistical Association*. 2001. Gibbs sampling methods for stick breaking priors; p. 161–173.
- Ishwaran H, Zarepour M. Markov chain Monte Carlo in approximate Dirichlet and beta two-parameter process hierarchical models. *Biometrika*. 2000; 87:371–390.

- Jansen I, Hens N, Molenberghs G, Aerts M, Verbeke G, Kenward MG. The nature of sensitivity in monotone missing not at random models. *Computational Statistics and Data Analysis*. 2006; 50:830–858.
- Kaplan D, George R. A study of the power associated with testing factor mean differences under violations of factorial invariance. *Structural Equation Modeling: A Multidisciplinary Journal*. 1995; 2(2):101–118.
- Karabotsos G, Walker SG. Coherent psychometric modelling with Bayesian nonparametrics. *British Journal of Mathematical and Statistical Psychology*. (in press).
- Kass R, Tierney L, Kadane J. Approximate methods for assessing influence and sensitivity in bayesian analysis. *Biometrika*. 1989; 76:663–674.
- Lavine M. Sensitivity in bayesian statistics: the prior and the likelihood. *Journal of the American Statistical Association*. 1991; 86:396–399.
- Lee, SY. Structural equation modeling: A Bayesian approach. Wiley & Sons Ltd.; West Sussex, England: 2007.
- Lee SY, Lu B, Song XY. Semiparametric Bayesian analysis of structural equation models with fixed covariates. *Statistics in Medicine*. 2007; 27(13):2341–2360.
- Lee SY, Tang NS. Analysis of nonlinear structural equation models with nonignorable missing covariates and ordered categorical data. *Statistica Sinica*. 2006; 16:1117–1141.
- Little, RJA., Rubin, DB. Statistical Analysis With Missing Data. Wiley; New York: 2002.
- Lu ZH, Chow SM, Loken E. Bayesian factor analysis as a variable-selection problem: Alternative priors and consequences. *Multivariate Behavioral Research*. 2016; 51(4):519–539. [PubMed: 27314566]
- Lütkepohl, H. Introduction to multiple time series analysis. 2nd. Springer-Verlag; New York: 2005.
- Millar RB, Stewart WS. Assessment of locally influential observations in bayesian models. *Bayesian Analysis*. 2007; 2:365–384.
- Molenaar PCM. A dynamic factor model for the analysis of multivariate time series. *Psychometrika*. 1985; 50:181–202.
- Molenaar, PCM. Dynamic factor analysis of psychophysiological signals. In: Jennings, JR, Ackles, PK., Coles, MGH., editors. *Advances in psychophysiology: A research annual*. Vol. 5. Jessica Kingsley Publishers; London, England: 1994. p. 229–302.
- Musher-Eizenman DR, Nesselroade JR, Schmitz B. Perceived control and academic performance: A comparison of high- and low-performing children on within-person change patterns. *International Journal of Behavioral Development*. 2002; 26:540–547.
- Muthén BO, Asparouhov T. Bayesian SEM: A more flexible representation of substantive theory. *Psychological Methods*. 2012; 17(1):313–335. [PubMed: 22962886]
- Navarro DJ, Griffiths TL, Steyvers M, Lee MD. Modeling individual differences using Dirichlet processes. *Journal of Mathematical Psychology*. 2006; 50:101–122.
- Nesselroade, JR., McArdle, JJ., Aggen, SH., Meyers, JM. Dynamic factor analysis models for representing process in multivariate time-series. In: Moskowitz, DS., Hershberger, SL., editors. *Modeling intraindividual variability with repeated measures data: Methods and applications*. Lawrence Erlbaum Associates; Mahwah, NJ: 2002. p. 235–265.
- Plummer M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*. 2003
- Rotnitzky A, Wypij D. A note on the bias of estimators with missing data. *Biometrics*. 1994; 50(4): 1163–1170. [PubMed: 7786998]
- Sbarra DA. Predicting the onset of emotional recovery following nonmarital relationship dissolution: Survival analyses of sadness and anger. *Personality and Social Psychology Bulletin*. 2006; 32:298–312. [PubMed: 16455858]
- Sethuraman J. A constructive definition of dirichlet priors. *Statistica Sinica*. 1994; 4:639–650.
- Shi X, Zhu H, Ibrahim J. Local influence for generalized linear models with missing covariates. *Biometrics*. 2009; 65:1164–1174. [PubMed: 19210738]
- Tanner M, Wong W. The calculation of posterior distributions by data augmentation (with discussion). *Journal of the American Statistical Association*. 1987; 82:528–550.

- Troxel AB, Lipsitz SR, Harrington DP. Marginal models for the analysis of longitudinal measurements with nonignorable non-monotone missing data. *Biometrika*. 1998; 85:661–672.
- Troxel AB, Ma G, Heitjan DF. An index of local sensitivity to nonignorability. *Statistica Sinica*. 2004; 14:1221–1237.
- van der Linde A. Local influence on posterior distributions under multiplicative modes of perturbation. *Bayesian Analysis*. 2007; 2:319–332.
- van Steen K, Molenberghs G, Verbeke G, Thijs H. A local influence approach to sensitivity analysis of incomplete longitudinal ordinal data. *Statistical Modelling*. 2001; 1(2):125–142.
- Verbeke G, Molenberghs G, Thijs H, Lasaffre E, Kenward MG. Sensitivity analysis for non-random dropout: A local influence approach. *Biometrics*. 2001; 57:43–50. [PubMed: 11252617]
- Weiss R. An approach to Bayesian sensitivity analysis. *Journal of the Royal Statistical Society Series B*. 1996; 58:739–750.
- Zautra AJ, Reich JW, Davis MC, Potter PT, Nicolson NA. The role of stressful events in the relationship between positive and negative affects: Evidence from field and experimental studies. *Journal of Personality*. 2000; 68:927–951. [PubMed: 11001154]
- Zhu H, Ibrahim JG, Chi YY, Tang N. Bayesian influence measures for joint models for longitudinal and survival data. *Biometrics*. 2012; 68(3):954–964. [PubMed: 22385010]
- Zhu H, Ibrahim JG, Lee SY, Zhang HP. Appropriate perturbation and influence measures in local influence. *Annals of Statistics*. 2007; 35(6):2565–2588.
- Zhu H, Ibrahim JG, Tang N. Bayesian sensitivity analysis of statistical models with missing data. *Statistica Sinica*. 2014; 24(2):871–896. [PubMed: 24753718]
- Zhu H, Ibrahim JG, Tang NS. Bayesian local influence analysis: A geometric approach. *Biometrika*. 2011; 98:307–323. [PubMed: 24453379]

Figure 1.

Plots of local influence measures showing (A) g_{ii} , (B) local influence measure based on the Bayes factor, FI_{BF, e_j} , (C) local influence measure based on ϕ -divergence, $SI_{D\phi, e_j}$, and (D) local influence measure based on the posterior means, $FI_{M_{d, e_j}}$ for Illustration II. The index values of 1–8 correspond to the perturbations in $\omega = (\omega_1, \dots, \omega_8)^T$.

Figure 2.

Panel (A) g_{ii} and index plots of the local influence measures: (B) FI_{BF, e_j} , (C) SI_{D, e_j} , and (D) FI_{M_d, e_j} for illustration IV. The indices 1, ..., 59 shown on the abscissa correspond to the 59 perturbation components in ω , namely, $\omega = (\omega_1, \dots, \omega_n, \omega_\mu, \omega_\zeta, \omega_\phi, \omega_{\tau1}, \omega_{\tau2}, \omega_y, \omega_{z1}, \omega_{z2}, \omega_a)^T$. Thus, the first 50 perturbations were perturbations to the 50 hypothetical subjects' data scores.

Figure 3. Panel (A) g_{ii} and index plots of the local influence measures: (B) FI_{BF, e_j} , (C) SI_{D, e_j} , and (D) FI_{M_d, e_j} for illustration V. The indices 1, ..., 6 shown in the abscissa correspond to the six perturbation components in $\omega = (\omega_z, \omega_{\mu_{z1}}, \omega_{\mu_{z2}}, \omega_y, \omega_\phi, \omega_\zeta)^T$.

Figure 4.

Index plots of local influence measures for the empirical example, showing (A) g_{ij} ; (B) FI_{BF,e_j} ; (C) SI_{D,ϕ,e_j} ; (D) FI_{M_d,e_j} computed using posterior mean functions involving all measurement parameters. The indices 1–188 correspond to the 188 perturbations in $\omega = (\omega_1, \dots, \omega_\eta, \omega_\mu, \omega_\zeta, \omega_\phi, \omega_{\tau 1}, \omega_{\tau 2}, \omega_{y1}, \omega_{y2}, \omega_{z1}, \omega_{z2}, \omega_\alpha, \omega_{b1}, \omega_{b2}, \omega_{b3}, \omega_{b4})^T$. The first 174 perturbations were perturbations to the participants' data scores.

Table 1

The First and Second-Order Local Influence Measures in Illustration III.

	FI_{BF, e_j}	$SI_{D\phi, e_j}$	FI_{M_d, e_j}
ω_1	0.026	0.164	0.192
ω_2	0.083	0.115	0.136
ω_3	115.337	231.889	255.997
ω_4	6.586	120.289	136.943

Table 2

Parameter Estimates from the Empirical Example Assuming NMAR and with Outlying Cases Removed, and Standardized Differences in Parameter Estimates between the Original and the Revised Models.

θ	Mean $\hat{\theta}$	SD $\hat{\theta}$	Std mean diff	θ	Mean $\hat{\theta}$	SD $\hat{\theta}$	Std mean diff
μ_1	0.102	0.031	-1.346	τ_{12}	-0.212	0.011	-0.021
μ_2	0.100	0.038	-0.975	τ_{13}	0.385	0.017	0.140
μ_3	0.102	0.034	-1.069	τ_{14}	0.791	0.021	0.054
μ_4	0.076	0.030	-0.957	τ_{15}	1.382	0.027	0.208
μ_5	0.068	0.027	<u>-2.168</u>	τ_{22}	-1.159	0.017	-0.035
μ_6	0.085	0.028	<u>-2.386</u>	τ_{23}	-0.461	0.017	0.245
μ_7	0.085	0.024	-1.360	τ_{24}	0.135	0.017	0.736
μ_8	0.078	0.021	-1.286	τ_{25}	0.767	0.019	0.897
λ_{21}	1.099	0.028	0.787	τ_{32}	-0.876	0.012	0.056
λ_{31}	1.055	0.030	0.886	τ_{33}	-0.305	0.014	0.318
λ_{41}	0.907	0.022	1.464	τ_{34}	0.222	0.015	0.640
λ_{62}	0.987	0.020	<u>-3.962</u>	τ_{35}	0.896	0.017	0.644
λ_{72}	0.814	0.021	-1.844	τ_{42}	-0.594	0.016	0.180
λ_{82}	0.666	0.015	<u>-3.943</u>	τ_{43}	0.214	0.018	0.740
ρ_0	-7.885	0.170	<u>-17.196</u>	τ_{44}	0.684	0.020	1.052
φ_1	2.331	0.083	<u>28.212</u>	τ_{45}	1.244	0.022	1.028
φ_2	-0.409	0.055	<u>-7.493</u>	τ_{52}	1.033	0.000	<u>7.901</u>
$\psi_{\zeta 11}$	0.083	0.005	-0.359	τ_{53}	1.954	0.000	<u>7.738</u>
$\psi_{\zeta 12}$	-0.093	0.005	-0.651	τ_{54}	2.541	0.000	<u>7.895</u>
$\psi_{\zeta 22}$	0.339	0.013	<u>9.941</u>	τ_{55}	2.942	0.000	<u>5.445</u>
Ψ_{e1}	0.419	0.012	1.144	τ_{62}	0.744	0.020	<u>2.168</u>
Ψ_{e2}	0.374	0.010	1.119	τ_{63}	1.606	0.033	<u>2.424</u>
Ψ_{e3}	0.415	0.010	0.659	τ_{64}	2.107	0.040	<u>2.344</u>
Ψ_{e4}	0.572	0.013	1.197	τ_{65}	2.591	0.039	1.977

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

θ	Mean $\hat{\theta}$	SD $\hat{\theta}$	Std mean diff	θ	Mean $\hat{\theta}$	SD $\hat{\theta}$	Std mean diff
ψ_{e5}	0.553	0.013	<u>7.187</u>	τ_{72}	0.928	0.024	<u>4.004</u>
ψ_{e6}	0.333	0.016	0.884	τ_{73}	1.889	0.038	<u>4.243</u>
ψ_{e7}	0.677	0.027	<u>3.715</u>	τ_{74}	2.411	0.044	<u>3.850</u>
ψ_{e8}	0.706	0.019	<u>3.145</u>	τ_{75}	2.881	0.039	<u>3.023</u>
				τ_{82}	0.242	0.015	1.851
				τ_{83}	1.059	0.022	<u>2.468</u>
				τ_{84}	1.575	0.025	<u>2.702</u>
				τ_{85}	2.042	0.026	<u>2.430</u>

Note: 95% HPD intervals for b_{11} , b_{22} , b_{12} , and b_{21} before the removal of outlying cases and under the assumption of MAR were given respectively by: [0.67, 1.05], [−0.33, 0.37], [−0.05, 0.27], and [−0.02, 0.96]. After the removal of outlying cases and with the NMAR missing data model, the 95% HPD intervals for b_{11} , b_{22} , b_{12} , and b_{21} were given respectively by: [0.46, 1.32], [−0.54, 0.48], [−0.22, 0.74] and [0.11, 1.06].