

A Bayesian Vector Autoregressive Model with Nonignorable Missingness in Dependent Variables and Covariates: Development, Evaluation, and Application to Family Processes

Linying Ji^{1,*}, Meng Chen^{1,*}, Zita Oravecz¹,
E. Mark Cummings², Zhao-Hua Lu³, Sy-Miin Chow¹

¹The Pennsylvania State University,

²University of Notre Dame,

³St. Jude Children's Research Hospital

Author Note

*These two authors contributed equally to the work.

Funding for this study was provided by NSF grant IGE-1806874, NIH grants R01GM105004 and U24AA027684, Penn State Quantitative Social Sciences Initiative and UL TR000127 from the National Center for Advancing Translational Sciences. Computations for Bayesian analyses were performed on the Pennsylvania State University's Institute for CyberScience Advanced CyberInfrastructure (ICS-ACI). Correspondence concerning this article can be addressed to Linying Ji, the Pennsylvania State University, 420 Biobehavioral Health Building, University Park, PA 16802 or by email to lj114@psu.edu.

Abstract

Intensive longitudinal designs involving repeated assessments of constructs often face the problems of nonignorable attrition and selected omission of responses on particular occasions. However, time series models, such as vector autoregressive (VAR) models, are often fit to these data without consideration of nonignorable missingness. We introduce a Bayesian model that simultaneously represents the over-time dependencies in multivariate, multiple-subject time series data via a VAR model, and possible ignorable and nonignorable missingness in the data. We provide software code for implementing this model with application to an empirical data set. Moreover, simulation results comparing the joint approach with two-step multiple imputation procedures are included to shed light on the relative strengths and weaknesses of these approaches in practical data analytic scenarios.

Keywords: Intensive longitudinal data, Bayesian vector autoregressive model, Multiple imputation, Nonignorable missing data

A Bayesian Vector Autoregressive Model with Nonignorable Missingness in Dependent Variables and Covariates: Development, Evaluation, and Application to Family Processes

Intensive longitudinal studies have gained popularity in the past two decades as a way of studying intra-individual change over time, inter-individual differences in intra-individual change, as well as other determinants of intra- and inter-individual variations ([Baltes and Nesselroade, 1979](#)). Such growth in popularity is facilitated in part by technological advances such as smartphone and wearable devices, which have provided enriched opportunities for intensive but relatively unobtrusive data collection in individuals' natural environment. Comparing with longitudinal panel data, these data are able to convey more nuanced information concerning change processes ([Bolger and Laurenceau, 2013](#); [Stone et al., 2008](#)).

One way to model these intensive longitudinal data (ILD) is through vector autoregressive (VAR) models, which can capture the temporal dependencies amongst multiple variables. VAR models have long been used in the econometric literature to forecast the performance of the stock market, macroeconomics trends, and policy change (for summaries see e.g., [Fomby et al., 2013](#); [Qin, 2011](#)). They are also gaining traction in recent years as a way of examining complex system dynamics in behavioral sciences, such as patterns of neural interactions in neuroimaging studies ([Ding et al., 2006](#)), intraindividual covariation of behaviors associated with personality models ([Hamaker et al., 2005](#)), daily patterns of emotional states and substance use among young adults in recovery ([Zheng et al., 2013](#)), dyadic interactions and coordinations between parent-child or partners ([Chow et al., 2010a](#); [Ram et al., 2014](#); [Thomas and Martin, 1976](#)), the different forms of concordance of a person's physiological responding to emotional stimuli ([Bulteel et al., 2014](#)), and network models of causal interplay between psychopathology symptoms ([Borsboom and Cramer, 2013](#); [Schmittmann et al., 2013](#)). In sum, these models are helpful in capturing more nuanced changes within each unit of analysis (e.g., within-person, within-dyad, and within-family dynamics). Despite the increased use of VAR models in

substantive applications, one ubiquitous challenge faced by many researchers is the lack of easily accessible methods of handling missingness in ILD, particularly in the presence of missing categorical covariates.

The current paper provides a review and illustrations of commonly adopted missing data handling techniques for VAR models. In particular, we include targeted comparisons of a recent two-stage, hybrid Bayesian-frequentist multiple imputation (MI) approach ([Ji et al., 2018](#)) to a novel single-stage, fully Bayesian approach proposed in this article. The illustrative example used throughout this article was motivated by an empirical study examining emotional dynamics of husband and wife after each conflict episodes, with the influence of children’s reaction to the conflicts, over a period of 15 days. A VAR model was applied to investigate how husband and wife’s emotional states were influenced by the emotional states of themselves, as well as their partner’s, after previous conflict episode. Two covariates on child reactions were also included in the model to reflect the child’s influence over parents’ emotional states. As detailed later, one big challenge in fitting a VAR model to this data set was the substantial amount of missingness in the child variables. This paper illustrates and investigates the extent to which appropriate inferential results can be obtained under different missing data handling approaches.

The present article has two unique contributions. In particular: (1) demonstrate a single-stage Bayesian modeling procedure to fit a VAR model with different missing data conditions with annotated sample modeling scripts; and (2) compare the performance of the single-stage Bayesian approach to a recent two-stage hybrid approach for handling missingness in ILD ([Ji et al., 2018](#)).

The remainder of this article is structured as follows. We first provide a review of common missing data handling methods in the literature and their applicability to ILD. We then outline the key features of the Bayesian missing data handling approach proposed in this study and how they differ compared to previous methods in the literature, including a two-step hybrid approach proposed by [Ji et al. \(2018\)](#), which combines a Bayesian approach

for imputation of missing covariates, and a frequentist approach for model estimation with imputed data. This is followed by an illustration of our proposed fully Bayesian approach, including sample coding syntax, using an empirical example. Finally, results from a Monte Carlo (MC) simulation study are presented to demonstrate the performance differences between the Bayesian approach and the two-step hybrid approach. We conclude with some recommendations for viable ways to handle potentially nonignorable missingness in ILD.

Missingness in Intensive Longitudinal Data and Common Approaches

We consider the value of a variable for one or multiple time points missing if it is unobserved or unreported. ILD is especially prone to missingness due to, for example, either study compliance issues or the nature of study design. Intensive assessments, especially those involving self-reports, require an ongoing commitment to remain in a study. The prolonged time span of these studies often entails increased participation burden over time, and in turn, higher likelihood of missingness.

Based on [Rubin \(1976\)](#), missing data mechanisms can be classified into three types: missing completely at random (MCAR), missing at random (MAR) and not missing at random (NMAR), which is also referred to as nonignorable missingness. If the mechanism is MCAR, the reason for missingness is independent of the questions investigated in the study (e.g., technical hiccup). MAR occurs when the probability of having missing data depends solely on some observed data (e.g., failure to report on weekends), but not on any unobserved information ([Fahrenberg and Myrtek, 2001](#)). Finally, we face nonignorable missingness when the reason for missingness is the unobserved missing data itself, meaning that the missingness mechanism relates to the question(s) we study (e.g., not giving reports on emotional experiences when being upset).

Typically, researchers make an assumption of the missing data mechanism, and then choose a method to handle it based on that assumption. It has been shown that wrong assumption of the missingness mechanism results in estimation problems, such as increased biases in point estimates ([Allison, 2003](#); [Jones, 1996](#)) and standard error estimates

(Glasser, 1964). Most contemporary work involving VAR models assumes the presence of MCAR or MAR. However, behavioral research with ILD involving self-report of constructs such as emotions are often subject to nonignorable missingness — for example, when the participants’ likelihood to report depends on their current emotions. If the data are NMAR, the missing data mechanism needs to be accounted for during the estimation process to yield unbiased parameter estimates.

Despite the proliferation of studies evaluating different missing data handling techniques in cross-sectional and longitudinal panel data with a limited number of measurement occasions (Allison, 1987; De Silva et al., 2017; Rubin, 1996; Schafer, 2001; Schafer and Graham, 2002; Sinharay et al., 2001), the impact of different kinds of missingness and strategies to handle missingness in multivariate, multi-subject time series data is less well studied. Full-information maximum likelihood (FIML) and MI are two most widely implemented approaches to handle missingness in the context of ILD, but they both have limitations. FIML operates by using only the available observed variables from each occasion and unit of analysis to compute the log-likelihood function for parameter and standard error estimation purposes. FIML is widely implemented in most software packages for VAR modeling, but it requires the missing values to be ignorable (i.e., MCAR or MAR), and the covariates in the model need to be fully observed.

In contrast to FIML, the MI approach requires the specification of a hypothetical model, termed the imputation model, to generate multiple “complete data sets” filled in with plausible values for the missing entries. This is followed by pooling of estimation results across all of the imputed data sets to yield a set of final parameter and standard error estimates. To better explicate this approach, we first define some notation as follows. Let \mathbf{Y} be an array of J dependent variables, \mathbf{y}_j ($j = 1, \dots, J$), for all individuals and time points, that can be partitioned into $\mathbf{Y} = \{\mathbf{Y}^{miss}, \mathbf{Y}^{obs}\}$. \mathbf{Y}_{-j} represents all other dependent variables except for the j th dependent variable. \mathbf{X} denotes an array of covariates for all individuals and time points. \mathbf{R} is a binary array summarizing the missing data patterns for

all the variables, with 0 indicating an observed entry and 1 indicating a missing entry. ϕ_j is the vector of parameters in the imputation model for \mathbf{y}_j^{miss} .

One approach to conduct MI, which is also the approach adopted in the hybrid method proposed by [Ji et al. \(2018\)](#), is multivariate imputation by chained equations (MICE), also referred to as full conditional specification ([van Buuren, 2012](#); [van Buuren and Groothuis-Oudshoorn, 2011](#)). With this approach, probable values of missing observations are generated in multivariate data on a variable-by-variable basis with a conditional probability model $P(\mathbf{Y}_j^{miss} | \mathbf{Y}_{-j}, \mathbf{X}, \mathbf{R}, \phi_j)$.¹ MICE has been widely adopted in cross-sectional and longitudinal panel studies, but its performance has been shown to be less satisfactory when used in intensive longitudinal settings ([Liu and Molenaar, 2014](#)).

To circumvent some known weaknesses of standard MICE techniques, [Ji et al. \(2018\)](#) proposed including lagged variables as predictors in the MICE imputation model, and compared different missing data handling techniques in the context of a VAR model. Two two-stage approaches were proposed: a full MI approach, in which all missing information were imputed before model fitting, and a partial MI approach, in which only missing covariates were multiply imputed, whereas missing data in the dependent variables were handled with FIML. Simulation results show that the two MI approaches outperformed simpler methods such as listwise deletion (LD). However, relatively large biases and poor coverage were still observed in some modeling parameters when mild misspecification was present in the imputation model. Given the two-stage nature of [Ji et al.](#)'s proposed approaches, the extent of sampling variability (i.e., standard error estimates) might have

¹ To start the procedure, all missing observations are filled in using random draws with replacement from all observed values. Then, for the first iteration, ϕ_1^1 is drawn from the distribution $P(\phi_1^1 | \mathbf{y}_1^{obs}, \mathbf{Y}_{-1}^0, \mathbf{X}^0, \mathbf{R})$. Missing values in \mathbf{y}_1 , \mathbf{y}_1^{miss} , are then filled in by drawing values from $P(\mathbf{y}_1 | \mathbf{y}_1^{obs}, \mathbf{Y}_{-1}^0, \mathbf{X}^0, \mathbf{R}, \phi_1^1)$, the posterior predictive distribution of \mathbf{y}_1 conditioned on \mathbf{y}_1^{obs} , \mathbf{Y}_{-1}^0 , \mathbf{X}^0 , \mathbf{R} , and ϕ_1^1 . Here, we use superscript k to denote data sets and parameter estimates from the k^{th} iteration, with $k = 0$ denoting the original data sets or initial starting values of the parameters. Similar procedure is subsequently performed to generate predicted values for \mathbf{y}_2^{miss} , only that imputed values for \mathbf{y}_1 from the previous step are used in the prediction process. The first iteration ends when missing observations are filled in for all variables. The procedure is repeated for K iterations to result in one set of data with imputed values. The whole procedure is repeated multiple times to generate multiple imputed data sets and correspondingly, multiple sets of parameter and standard error estimates for subsequent pooling ([van Buuren, 2012](#)).

been underestimated as well. In this study, we extend the study by [Ji et al.](#) by comparing the authors' recommended two-stage approach with a Bayesian single-stage approach, and under conditions with a broader array of missing data percentages and degrees of non-ignorability in missingness ([Graham, 2012](#)). Since better parameter point estimates for some of the parameters were obtained with the partial MI approach in the previous simulation study, we will only consider the partial MI approach in this paper.

A Bayesian Approach to Handle Missing Data

The limitations of the FIML and MI approaches in handling particular missing data mechanisms and underestimation of sampling variability can be readily circumvented with the proposed Bayesian approach. In what follows we highlight the key properties of the Bayesian approach to joint modeling of the missingness mechanism and the multivariate processes of interest via a VAR model. Specifically, we specify a joint probabilistic model for the *full data*, with explicit assumptions about the mechanism that have given rise to the missingness in the data. This is done by building submodels that describe the probabilities that the covariates and/or endogenous variables are missing in the data set. These missing data submodels are embedded within a larger model that also includes models of interest (in our case, a VAR model) that specify the time evolution of the intensively measured variables and their latent variable counterparts – whether they are observed or missing for particular occasions. In cases where covariates are not fully observed, a model that describes the distributions of the covariates is also necessary as a part of this larger model. We will refer to this larger model as the “full-data model” for the remainder of this article.

[Daniels and Hogan \(2008\)](#) illustrated three different ways to factorize the full-data distribution, namely, as selection models ([Diggle and Kenward, 1994](#); [Heckman, 1979](#)), mixture models ([Hogan and Laird, 1997](#); [Little, 1993,9](#)) and shared parameter models ([Henderson et al., 2000](#); [Wu and Carroll, 1988](#)). The full-data model can be specified as $P(\mathbf{Y}, \mathbf{R} | \mathbf{X}, \boldsymbol{\omega})$, where \mathbf{Y} , \mathbf{X} and \mathbf{R} are as defined earlier in the introduction to MI in the previous section, and $\boldsymbol{\omega}$ is a collection of model parameters. Under the selection model, the

full-data model can be factorized as follows:

$$P(\mathbf{Y}, \mathbf{R}|\mathbf{X}, \boldsymbol{\omega}) = P(\mathbf{Y}|\mathbf{X}, \boldsymbol{\theta})P(\mathbf{R}|\mathbf{Y}, \mathbf{X}, \boldsymbol{\phi}), \quad (1)$$

where $P(\mathbf{Y}|\mathbf{X}, \boldsymbol{\theta})$ represents the model for dependent variables (i.e., model of interest), and $P(\mathbf{R}|\mathbf{Y}, \mathbf{X}, \boldsymbol{\phi})$ represents the missing data mechanism. For example, [Diggle and Kenward \(1994\)](#) proposed a model for continuous longitudinal data with NMAR drop-out, where a multivariate linear model was used as the model of interest, and the drop-out process was modeled using a logistic regression model, where probability of missing an observation depended on observed response history and the current value of dependent variable, which might be missing. The mixture and shared parameter modeling approaches capitalize on different assumptions and hence, different specifications for the conditional relations among elements of the full-data distribution². In this paper, we illustrated the use of selection model to handle missing data in intensive longitudinal studies because of the following benefits: First, compared with the mixture models, the selection model is more feasible in the context of ILD. This is because the mixture model approach involves conditioning of \mathbf{Y} over each possible combination of missing data patterns. With ILD, there could be too many possible missing data patterns, and relatively few observations within each pattern. In addition, slight differences in patterns may not reflect meaningful individual differences at all. Second, unlike the shared parameter approach, the model of substantive interest, $P(\mathbf{Y}|\mathbf{X}, \boldsymbol{\omega})$, as well as the missing data mechanism are directly specified in the selection model approach, which is very intuitive for researchers. Finally, the selection model is

² The mixture models factor the full-data model as:

$$P(\mathbf{Y}, \mathbf{R}|\mathbf{X}, \boldsymbol{\omega}) = P(\mathbf{Y}|\mathbf{R}, \mathbf{X}, \boldsymbol{\omega})P(\mathbf{R}|\mathbf{X}, \boldsymbol{\omega}),$$

With this model, the relations between \mathbf{Y} and \mathbf{X} are conditioned on different missing data patterns. In contrast, the shared parameter approach assumes a multilevel structure and models random effects \mathbf{b} jointly with \mathbf{Y} and \mathbf{R} with following general model:

$$P(\mathbf{Y}, \mathbf{R}|\mathbf{X}, \boldsymbol{\omega}) = \int P(\mathbf{Y}, \mathbf{R}, \mathbf{b}|\mathbf{X}, \boldsymbol{\omega})d\mathbf{b}.$$

relatively easy to estimate, as compared with mixture models, which may have identification issues with increased dimension of \mathbf{Y} , and shared parameter models, which involve integration over the shared parameters.

Given the full-data model, parameter estimation and inference in the Bayesian statistical framework are based on the posterior distributions of the model parameters and missing data. Posteriors are derived based on priors of model parameters and the likelihood of the observed data, combined via Bayes' rule. The missing data are treated as an unknown quantity, similar to the unknown model parameters. The unknown quantities are explored through Markov chain Monte Carlo (MCMC) techniques according to the posterior distribution. If prior information on the possible values of the parameters is not available, uninformative priors are set. However, when fitting a VAR model, we suggest to use weakly informative priors for dynamic parameters that impose, whenever possible and appropriate, the stationarity assumptions — namely, the assumption that the joint distribution of the time series variables is time invariant ([Lütkepohl, 2005](#)).

The proposed Bayesian implementation has several advantages over the MI approaches. First, unlike the two-stage MI approaches (e.g. [Ji et al., 2018](#)), the proposed Bayesian implementation estimates the underlying dynamics and missingness mechanism in a single step ([Daniels and Hogan, 2008](#)). That is, in the MI approaches, a MI step is first used to impute values for the missing observations in \mathbf{Y} and \mathbf{X} , typically based on some overparameterized generalizations of $P(\mathbf{Y}|\mathbf{X}, \boldsymbol{\omega}, \mathbf{R})$ that often include some additional *auxiliary variables* in \mathbf{X} that help explain \mathbf{R} , but not necessarily the dynamics of \mathbf{Y} directly. Once the imputed values of \mathbf{Y} and \mathbf{X} are available, only $P(\mathbf{Y}|\mathbf{X}, \boldsymbol{\omega})$ is modeled directly in the second stage as if all \mathbf{Y} and \mathbf{X} values were fully observed. With the Bayesian selection modeling approach, both $P(\mathbf{Y}|\mathbf{X}, \boldsymbol{\omega})$ and $P(\mathbf{R}|\mathbf{Y}, \mathbf{X}, \boldsymbol{\phi})$ are estimated simultaneously over long chains of iterative MCMC updates. This is to say that with this approach, we simultaneously estimate the missing data modeling parameters and the VAR modeling parameters. The missing data estimates can help researchers learn about the

underlying missingness mechanism. Second, compared with MI approaches, the Bayesian approach allows more flexibility in missing data model specification. Different missing data models may be specified for missingness in different variables, and different missing data mechanisms (i.e., MCAR, MAR, and NMAR) can be fully reflected in the model specification. Researchers may also incorporate any theory-driven models of missingness into the larger model. Third, in terms of standard error estimates, MI approaches only incorporate a limited number of imputations, and this number might underestimate the uncertainty in the imputation procedure. With Bayesian methods, the uncertainty about the imputed data is fully propagated into all subparts of the estimation process, over numerous iterations. Thus, standard errors and related quantities for inferential purposes reflect the multiple sources of uncertainty from the data. Fourth, Bayesian methods allow estimation of more complex models and data structure, such as multilevel models with random effects ([Wang and McArdle, 2008](#)), and data with nested structure. Previous studies have demonstrated the strengths of the Bayesian missing data modeling approach with different models, such as multiple regression models with observational studies with NMAR missing data ([Mason et al., 2010](#)), analysis of randomized clinical trials with drop-outs ([Scharfstein et al., 2003](#); [Wang et al., 2010](#)), path analysis model ([Gajewski et al., 2006](#)) where the outcome variable and the mediating variable follow Poisson distributions, survival analysis model ([Hemming and Hutton, 2012](#)), hierarchical models for network meta-analysis ([Zhang et al., 2015](#)), and nonparametric statistical learning models with Bayesian Additive Regression Trees ([Kapelner and Bleich, 2015](#)). We add to this literature by studying the performance of the Bayesian missing data approach in modeling multivariate, multi-subject intensive time series data under different missing data mechanisms.

Empirical Illustration

Motivating Model and Methodological Challenges

The proposed Bayesian VAR model was inspired by a previously published study that explored the emotional dynamics of interparental conflicts ([Schermerhorn et al., 2010](#)). Researchers collected data on emotional states at the end of conflicts with influences from child emotions and behaviors during conflicts from 111 cohabiting couples with a child (child ages range from 8 to 16 years), over 15 days. This study was an event-contingent design where parents responded only during or shortly after a conflict. The parents were asked to record their own emotional states as well as their children's emotional states and behaviors associated with particular conflicts.

We proposed the following VAR model to capture the time-dynamics of the observed data:

$$\begin{bmatrix} w_{nt} \\ h_{nt} \end{bmatrix} = \begin{bmatrix} a_{w \rightarrow w} & b_{h \rightarrow w} \\ b_{w \rightarrow h} & a_{h \rightarrow h} \end{bmatrix} \begin{bmatrix} w_{n,t-1} \\ h_{n,t-1} \end{bmatrix} + \begin{bmatrix} c_{x_1 \rightarrow w} & d_{x_2 \rightarrow w} \\ c_{x_1 \rightarrow h} & d_{x_2 \rightarrow h} \end{bmatrix} \begin{bmatrix} x_{1nt} \\ x_{2nt} \end{bmatrix} + \begin{bmatrix} \epsilon_{w,n,t} \\ \epsilon_{h,n,t} \end{bmatrix} \quad (2)$$

$$\epsilon_{w,n,t} \sim \mathcal{N}(0, \sigma_{\epsilon_w}^2), \quad \epsilon_{h,n,t} \sim \mathcal{N}(0, \sigma_{\epsilon_h}^2)$$

where w_{nt} and h_{nt} represented conflict resolution ratings of wife and husband, respectively, from family n at the end of the t^{th} conflict ($n = 1, \dots, N$; $t = 1, \dots, T_n$). Given the event-contingency, the total number of conflicts (T) was different for every family ($min = 11, max = 69, mean = 25.90$). The terms $\epsilon_{w,n,t}$ and $\epsilon_{h,n,t}$ were the residuals for wife and husband not accounted for by the hypothesized model, assumed to be normally distributed with zero means, standard deviations σ_{ϵ_w} and σ_{ϵ_h} , respectively.

The hypothesized model was a VAR model of order 1 in which the dependent variables at the current time point were predicted by the dependent variables at the immediately preceding time point (i.e., a lag of 1). In the present context, the emotional states and conflict resolution behavior of each spouse at the end of the t^{th} conflict were posited to be influenced by their own emotional states and resolution behavior at the $t - 1^{th}$ conflict, the strengths of which were captured in the auto-regression parameters, $a_{w \rightarrow w}$ and $a_{h \rightarrow h}$. In

addition, each person’s previous emotional states and resolution behavior at the previous conflict were also assumed to affect the partner’s emotional state and resolution behavior at the t^{th} conflict, as governed by the cross regression parameters, $b_{h \rightarrow w}$ and $b_{w \rightarrow h}$.

Two covariates were included in the dynamic model. The covariate, x_1 , was an aggregate measure of the child’s negativity, as averaged across actions such as anger, sadness, fear, as well as misbehaving, yelling at the parents, and aggression. The other covariate, x_2 , represents a child aggregate score on agentic behavior in family i , which includes actions such as helping out, taking sides, comforting the parents, and trying to make peace. Each one of the dependent variables and the covariates were standardized by family over time prior to model fitting to ease prior selection and to remove some of the pre-existing interindividual differences in process noise variances and VAR dynamics — features not accommodated by our hypothesized model.

In this empirical study, a large portion (67% for all child-related variables) of the child-related covariates was missing (Schmerhorn et al., 2010). In order to handle the missingness, the authors previously recoded the child-related covariates from sum scores (ranging from 0 to 10) into dummy variable such that a child’s value on each covariate was coded as 0 both when the child did not display the behavior during the conflict, and when the child was “missing”; each of the two covariates was coded as 1 when the child showed any level of that behavior. This coding scheme has three primary drawbacks: (1) the data blur levels of a child’s influence with the presence or absence of the data, obscuring the ability to make meaning of the data, (2) it discounts potential effects of different levels of the child-variables on dynamics at the family level, and (3) this data mechanism may be inappropriate as both the dependent and child-related covariates may be NMAR (e.g., the couples might be especially careful in ensuring that the child was absent when they anticipated discussing highly stress-provoking topics). In order to account for some of the aforementioned issues, Ji et al. (2018) applied a MI method to handle missing covariates in this data set. With this method, the authors were able to preserve information on the

magnitudes of child variables, since the variables were not dichotomously recoded.

However, the plausible missing data mechanism were still unknown and NMAR missing effects might not be fully accounted for.

Bayesian Joint VAR and NMAR Modeling Example with JAGS

We utilized the Bayesian approach to model the dynamics of the process and the missing data mechanisms simultaneously. With this approach, we can compare joint models for the dynamics data and hypothesized missingness mechanisms. Bayesian model fitting was performed in JAGS ([Plummer et al., 2003](#)), interfaced with R ([R Core Team, 2016](#)) for data formatting and result summaries through R package rjags ([Plummer, 2016](#)). Details of the model fitting procedures are presented below.

In the empirical data set, T_n corresponds to the total number of conflicts for the n th family, with T_{max} representing the maximum number of conflicts across all families. We structure the data as:

Dependent variables $\mathbf{Y}_1, \mathbf{Y}_2$: $N \times T_{max}$ matrices

Covariates $\mathbf{X}_1, \mathbf{X}_2$: $N \times T_{max}$ matrices

Missing data patterns (dummies) $\mathbf{R}_{Y_1}, \mathbf{R}_{Y_2}, \mathbf{R}_{X_1}, \mathbf{R}_{X_2}$: $N \times T_{max}$ matrices

Time series length for the n^{th} family, T_n : T[n]

For the Bayesian approach, the full-data model – including the VAR model for the dependent variables, the model for covariates, and the missingness model for both dependent variables and covariates – is specified in a text script to be read into JAGS. Since we have longitudinal data for multiple families and multiple observations for each family, the model follows the basic structure below:

```
model{
  for (n in 1:N){
    y1[n,1] ~ dnorm(0, 0.01)
    y2[n,1] ~ dnorm(0, 0.01)
```

```

x1[n,1] ~ dnorm(0, 0.01)
x2[n,1] ~ dnorm(0, 0.01)

for (t in 2:T[n]){
  <VAR model for dependent variables>
  <Model for covariates>
  <Model for missingness mechanisms>
} # end of t loop

} # end of n loop
<Priors>
}

```

Two for-loops are shown in the script above: one that loops through every family (the n loop) and one through every observational time point (the t loop), the maximum number of which is family-specific ($T[n]$). The first four lines inside of the n loop specify the distributions for the first observations. In this example we use a normal distribution with a large variance (100, or equivalently, a small precision of 0.01 as indicated in the script) for all four observations, so that the distribution itself does not add very specific information about the initial values of the variables at time 1, yet captures a reasonable range of starting values. One thing to note is that the `dnorm` function in JAGS parametrizes the normal distribution with its mean and precision, instead of variance (where $\text{precision} = \frac{1}{\text{variance}}$). The t loop contains three submodels: a VAR model for the covariates, a model for covariates, and a model for missingness mechanisms. We will introduce each of the models and the corresponding JAGS syntax in the following paragraphs, and a complete script for model fitting is included in Appendix C. The model for covariates is only necessary if the covariates are not fully observed.

VAR Model for Dependent Variables. We first write out the model for dependent variables, namely, the VAR model with covariates (based on Equation 2) as

$$\begin{aligned} y1[n,t] &\sim \text{dnorm}(\text{mu1}[n,t], \text{tau1}) \\ y2[n,t] &\sim \text{dnorm}(\text{mu2}[n,t], \text{tau2}) \\ \text{mu1}[n,t] &<- a*y1[n,t-1] + b*y2[n,t-1] + c*x1[n,t] + d*x2[n,t] \\ \text{mu2}[n,t] &<- b1*y1[n,t-1] + a1*y2[n,t-1] + c1*x1[n,t] + d1*x2[n,t] \end{aligned}$$

Model for Covariates. A model for covariates is also necessary since the covariates also contain missingness. In our illustrative example, both the covariates are measured using a numeric response scale. Thus, we assume that each covariate follow a normal distribution with:

$$x_{1nt} = \beta_1 + \beta_2 p_{wnt} + \beta_3 p_{hnt} + \epsilon_{x1,nt}, \quad \epsilon_{x1,nt} \sim \mathcal{N}(0, \sigma_{\epsilon_{x1}}^2) \quad (3)$$

$$x_{2nt} = \psi_1 + \psi_2 x_{1nt} + \psi_3 p_{wnt} + \psi_4 p_{hnt} + \psi_5 age_n + \epsilon_{x2,nt}, \quad \epsilon_{x2,nt} \sim \mathcal{N}(0, \sigma_{\epsilon_{x2}}^2). \quad (4)$$

Fully observed post-conflict positivity levels of husbands' (p_h) and wives' (p_w), both $N \times T_{max}$ matrices, are entered into the model for covariates above and the model for missingness mechanisms later. Since we do not have specific assumptions of mother's and father's positivity levels influencing children differently, we constrain p_w 's and p_h 's influence on x_1 and x_2 to be the same, as indicated by having the same parameter coefficient in Equations 3 and 4 (β_2 and ψ_3). This approach can be seen as being equivalent to taking the average levels between mother's and father's positivity scores as a predictor. age is specified as a vector of length N with the n th element being the age of the child in family n . These three variables (p_h , p_w , and age), albeit not variables of interest in the VAR model, are speculated to be related to the covariates and probabilities of missingness, and are thus included to improve the performance of the full-data model.

As with the model for the dependent variables, the model shown in Equations 4 is but one plausible model for the covariates. This model reflects our hypothesized relations among the measured covariates and other fully observed auxillary variables in the current

data set. That is, we assume that x_1 (the child’s negativity) would be affected by the parents’ positivity post-conflict. In addition, we assume that x_2 (the child’s agentic behaviors) is related to parents’ positivity, the child’s post-conflict negativity (x_1 ; e.g. [Davies and Cummings, 1994](#); [Emery, 1989](#)), and the child’s age. The age-based postulate is motivated by the theoretical expectation that agentic behaviors require certain levels of executive functioning that develop with age (e.g. [Goeke-Morey et al., 2013](#); [Grych and Fincham, 1990](#)).

This model for the covariates can be specified in JAGS script as follows:

```
x1[n,t] ~ dnorm(mux1[n,t],tau_x1)
x2[n,t] ~ dnorm(mux2[n,t],tau_x2)
mux1[n,t] <- beta[1] + beta[2]*pw[n,t] + beta[2]*ph[n,t]
mux2[n,t] <- psi[1]+psi[2]*x1[n,t]+psi[3]*pw[n,t]+
              psi[3]*ph[n,t]+psi[4]*age[n] .
```

In order to explore the sensitivity of modeling results to the model for covariates, we also considered an alternative, unconditional model for the covariates. In this alternative model, x_1 did not depend on any variable observed in the data set (Appendix [B](#)).

Model for Missingness Mechanisms. Next, we specify the missingness mechanisms for all the variables. All the missingness indicators ($r_{y_{1nt}}, r_{y_{2nt}}, r_{x_{1nt}}, r_{x_{2nt}}$) are entered into JAGS as observed data (1 = missing, 0 = observed). For this empirical data set, we assume that missingness in all of the dependent variables and covariates is associated with the post-conflict positivity scores of husbands and wives (p_h and p_w), as the couples’ emotional states post-conflict might affect their willingness to report. In addition, we assume that missingness in both of the covariates is associated with children’s agentic behaviors (x_2), as children’s tendency of such behaviors may affect their presence on scene.

Thus, we specify the missingness mechanisms as:

$$P(r_{y_{1nt}} = 1 | p_{w_{nt}}, p_{h_{nt}}) = \text{logit}^{-1}(\phi_{int_{y_1}} + \phi_{p \rightarrow y} p_{w_{nt}} + \phi_{p \rightarrow y} p_{h_{nt}}) \quad (5)$$

$$P(r_{y_{2nt}} = 1 | p_{w_{nt}}, p_{h_{nt}}) = \text{logit}^{-1}(\phi_{int_{y_2}} + \phi_{p \rightarrow y} p_{w_{nt}} + \phi_{p \rightarrow y} p_{h_{nt}}) \quad (6)$$

$$P(r_{x_{1nt}} = 1 | x_{2nt}, p_{w_{nt}}, p_{h_{nt}}) = \text{logit}^{-1}(\phi_{int_{x_1}} + \phi_x x_{2nt} + \phi_{p \rightarrow x} p_{w_{nt}} + \phi_{p \rightarrow x} p_{h_{nt}}) \quad (7)$$

$$P(r_{x_{2nt}} = 1 | x_{2nt}, p_{w_{nt}}, p_{h_{nt}}) = \text{logit}^{-1}(\phi_{int_{x_2}} + \phi_x x_{2nt} + \phi_{p \rightarrow x} p_{w_{nt}} + \phi_{p \rightarrow x} p_{h_{nt}}). \quad (8)$$

In the above equations, logit^{-1} represents the inverse-logit (logistic) function. This model for missingness mechanisms is a mixture of MAR missing models for the dependent variables, and NMAR missing models for the covariate variables. That is, the missing data model for the dependent variables consists only of post-conflict positivity scores of husbands and wives, which are fully observed, whereas the missing data model for the covariates contains x_2 , which is partially missing. In addition to this mixed missing mechanisms model, we also considered three other alternative missing mechanisms models, including MCAR, MAR, and NMAR models for all dependent variables and covariates respectively. With MCAR model, for all dependent and covariate variables, the missingness were modeled by an intercept, assuming missingness in all four variables were completely random and did not depend on any other variables. Under the MAR condition, missingness of all variables were predicted by the two fully observed variables, the positivity scores of husbands and wives. In the NMAR model, the missingness of the variables depended on the value of the variables themselves, in addition to the fully observed positivity scores of the couples. Model specification details for the alternative models are presented in Appendix [A](#).

Below is the corresponding script for specifying missingness mechanism for x_1 , child's negativity, as an example:

```
nmarlogitx1[n,t]<-phix1[1]+phix1[2]*x2[n,t]+phix1[3]*
pw[n,t]+phix1[3]*ph[n,t]
nmarprx1[n,t]<-exp(nmarlogitx1[n,t])/(1+exp(nmarlogitx1[n,t]))
```

$$R_{x1}[n, t] \sim \text{dbern}(\text{nmarprx1}[n, t])$$

The first line in the JAGS code specifies the linear combination in Equation 7 delineating the log odds of observing missingness in x_1 as related to other fully observed or partially missing variables. The second line transforms the linear combination through the logistic function into a missing data probability for family n at time t ($\text{nmarprx1}[n, t] = P(r_{x1_{nt}} = 1 | x_{2_{nt}}, p_{w_{nt}}, p_{h_{nt}})$). Finally, the third line of code specifies the missing data indicator in $r_{x1_{nt}}$ as Bernoulli distributed with a probability given by

$$P(r_{x1_{nt}} = 1 | x_{2_{nt}}, p_{w_{nt}}, p_{h_{nt}}).$$

Priors and Model Selections. In the present illustration, we use weakly informative priors (normal distributions with means at 0 and variances of 10) for parameters in the VAR model (i.e. a , $a1$, b , $b1$, c , $c1$, d , $d1$). Even though these priors may seem informative, the ranges of possible parameter values under these priors are still notably more diffuse than the values that would be expected to arise under the stationarity constraint imposed in generating the simulated data. The rest of the prior settings can be considered non-informative: we use a gamma prior with both parameters set to 0.001 for the inverse of the variance parameters (e.g. tau1), and a normal prior centered on 0 with variance 100 for all other parameters. Code for prior specification is included in the JAGS Script under Appendix C.

Given the complexity of the complete data models, we utilized the following setting in sampling: 2 chains, 50000 steps of adaptation, 10000 steps of burn-in, and 50000 samples per chain. To help with model convergence, we linearly interpolated values for the missing data based on the observed values immediately before and after the missing locations, and entered the interpolated values into JAGS as starting values for the MCMC sampler at these missing locations.

Based on previous simulation results (see, e.g., [Lu et al., 2017](#)), we used the Bayesian Information Criterion (BIC) as a model selection measure in the empirical illustration to select the models described above from several possible candidate models. We obtained

BIC from the MCMC estimation as:

$$BIC \approx \bar{D} + p \times \log(S) \quad (9)$$

In the above equation, \bar{D} is the posterior mean of the deviance, defined as:

$$\bar{D} = - \sum_{n=1}^N \sum_{t=1}^{T_n} \left(\log(f_Y) + \log(f_X) + \log(f_R) \right), \quad (10)$$

In Equation 9, p is the number of parameters in a given model, and S represents the sample size, which is taken to be the sum of all T 's across families in the present study.

Other terms in Equation 10 are defined as follows: $f_Y = f_Y(\mathbf{y}_{nt} | \mathbf{y}_{n,t-1}, \mathbf{x}_{n,t-1}, \boldsymbol{\theta})$, $f_X = f_X(\mathbf{x}_{nt} | \mathbf{v}_{nt}, \boldsymbol{\theta})$, $f_R = f_R(\mathbf{r}_{nt} | \mathbf{y}_{n,t}, \mathbf{x}_{n,t}, \boldsymbol{\theta})$, where \mathbf{y}_{nt} , \mathbf{x}_{nt} , \mathbf{v}_{nt} , \mathbf{r}_{nt} represent vectors of dependent variables $[w_{nt}, h_{nt}]$, covariates $[x_{1nt}, x_{2nt}]$, auxiliary variables $[p_{w_{nt}}, p_{h_{nt}}]$, and missingness indicators $[r_{w_{nt}}, r_{h_{nt}}, r_{x_{1nt}}, r_{x_{2nt}}]$ respectively, and f_Y , f_X , f_R are the conditional joint densities of \mathbf{y}_{nt} , \mathbf{x}_{nt} , and \mathbf{r}_{nt} . Note that our definition of BIC was similar to the standard definition for BIC adopted within the frequentist framework in that we did not add the number of missing data values as unknown parameters in the calculation of p .

We used the BIC, for instance, to compare our hypothesized model for the covariates in Equations 3 – 4 to the alternative, unconditional model for the covariates (Appendix B). Our hypothesized covariate model had a lower BIC value and was thus preferred over the alternative model. In addition, inspection of the parameter estimates from both models suggested that the key substantive conclusions obtained with the alternative covariate model remained the same compared to our original hypothesized model (see Appendix B). This indicated that the substantive results were not sensitive to such changes in the covariate model. We also used the BIC to select among several possible models for the missingness mechanisms. We compared the BIC values across the four proposed possible missingness mechanism models and selected the mixed missingness mechanisms model as the best-fitting model based on its smallest BIC value (see Appendix A for details).

In order to compare the estimation results from the proposed Bayesian method and the two-step partial MI method, the following procedures were performed for the partial MI

method. We first imputed multiple sets of plausible values for the missing observations, using all information available and the lagged dependent variables. Given the relative large percentage of missing data in the covariates, 10 imputations were performed, generating 10 sets of imputed data. The VAR model for dependent variables was then fitted using the imputed data sets, resulting in 10 set of parameter estimates. Finally, parameter estimates obtained from different imputed data sets were pooled. Detailed procedure for applying the two-step MI method and example syntax are available in [Ji et al. \(2018\)](#).

Empirical Results

Parameter estimates from both approaches are shown in Table 1. None of the dynamic parameters (neither the autoregressive nor the cross-regressive ones) was different from 0 (defined as credible/confidence intervals not containing 0). Results from the Bayesian approach suggested conflict resolution scores for both husbands and wives were negatively associated with children’s negativity scores and positively associated with children’s agentic behaviors ($c_{x_1 \rightarrow w} = -0.50$, $c_{x_1 \rightarrow h} = -0.51$, $d_{x_2 \rightarrow w} = 0.06$, $d_{x_2 \rightarrow h} = 0.06$). These are in accordance with previous literature on children’s influence towards marital conflict ([Schmerhorn et al., 2010](#)).

However, it is worth noting the differences in estimated parameter values between approaches. Results from the two-step partial MI approach, though also suggesting a negative association between children’s negativity scores and conflict resolutions, showed much smaller effect sizes on parameters $c_{x_1 \rightarrow w}$ and $c_{x_1 \rightarrow h}$ (-0.5 vs. -0.1). The effects of children’s agentic behaviors on conflict resolution were not significant under the partial MI approach, and the process noise variances were also estimated to be larger compared to the Bayesian approach. We speculate these differences to be due to the missingness mechanism being partially NMAR (as suggested by our hypothesized missingness mechanism model outperforming the MAR mechanism model in terms of BIC). This demonstrates the possible different conclusions drawn from an analysis if NMAR mechanism is not addressed, even when partial MI is used.

To check whether the parameter estimates from the Bayesian approach were sensitive to prior distributions, we performed post-hoc sensitivity check following the procedures described in Lee (2007) separately for the VAR model, the covariate model, and the missing data parameters. The sensitivity check involved altering the prior distributions in three ways and rerunning the analysis. The variances for the prior distributions were set to 1 for all parameters as opposed to 10 and 100 in the previous analysis, and the means were set to (1) estimated parameter values from the previous analysis based on the vague priors, (2) halves of the estimated parameter values and (3) two times the estimated parameter values. Table 2 shows the estimated values from all 9 sensitivity checks and most values remained unchanged from those resulted from the previous analyses. The only exception was the intercept parameters $(\phi_{int_{y_1}}, \phi_{int_{y_2}}, \phi_{int_{x_1}}, \phi_{int_{x_2}})$ and the NMAR parameter (ϕ_x) when the means of the missingness model parameter prior distributions were changed to twice their estimated values. Even though these differences may seem large on a logit scale, the differences in estimated values on a probability scale were much less notable. This suggests that the missing model parameters were relatively sensitive to the choices of prior distributions. This may due in part to the difficulties in obtaining sufficient effective sample sizes for these parameters despite our use of very long chains.

This empirical illustration demonstrated the utility of the Bayesian approach with data from an existing empirical study. Estimation results with Bayesian method were in general consistent with other approaches, such as two-step partial MI. However, larger covariate effects, and smaller noise variances were noted with the Bayesian method. Observing the estimated missing data model parameters with the Bayesian method, we can also infer that the missing data mechanism for the covariates was likely NMAR.

Simulation Study: Design and Results

We conducted an MC simulation study to evaluate the performance of different missing data handling methods under different missing data conditions. Factors that we manipulated include: missing data mechanism (MCAR, MAR, NMAR), strength of NMAR

dependency in the NMAR conditions, and missing data percentages. Our key goal was to clarify the performance of the fully Bayesian joint modeling approach relative to other missing data handling approaches, particularly the two-step approach evaluated previously by [Ji et al. \(2018\)](#).

For each condition, we conducted 100 MC replications using the bivariate VAR model in Equation 2. Missingness was imposed on each of the complete data sets using one of three possible missing data models: MCAR, MAR, and NMAR. The NMAR missing data model was the most general model because it included NMAR components, MAR components, as well as MCAR components (details of the missing data models to follow). In order to evaluate the influence of percentage of missing data and the strength of the NMAR component, we crossed the NMAR strength (low, high) with two possible (low, high) missingness percentages, resulting in four NMAR conditions, namely NMARL3 (low dependency NMAR with 30% of missing data), NMARL5 (low dependency NMAR with 50% of missing data), NMARH3 (high dependency NMAR with 30% of missing data), and NMARH5 (high dependency NMAR with 50% of missing data). The bivariate VAR model was fit to each of the simulated data sets using different missing data handling approaches, including the LD method, two-step partial MI method, and Bayesian approach. Next, we describe the data generating models, missing data settings, and model-fitting procedures in detail.

The Data Generating Model

The dynamical process in the simulation study was based on the VAR model introduced in the motivating example (Equation 2) and used in the Empirical Illustration section. The sample size configuration was chosen to mirror characteristics of the empirical data set. We simulated data for $N = 100$ couples. The total number of observations (T) for two people within each dyad was assumed to be the same, and T per dyad followed a negative binomial distribution of number of successes before five failures with a failing probability of 0.167. Figure 1 offers a comparison of the empirical and simulated Ts. The

observed simulated distribution was derived from simulated Ts in all 100 MC replications. We can see from the figure that the distribution of T of the simulated datasets mimics the distribution of the observed T of the empirical data set.

The true parameter values of the dynamic model used in the simulation were set as follows: $a_{w \rightarrow w}=0.4$, $a_{h \rightarrow h}=0.3$, $b_{h \rightarrow w}=-0.3$, $b_{w \rightarrow h}=-0.2$, $c_{x_1 \rightarrow w}=0.4$, $c_{x_1 \rightarrow h}=0.3$, $d_{x_2 \rightarrow w}=0.5$, $d_{x_2 \rightarrow h}=0.4$, and $\sigma_{\epsilon_w}^2 = \sigma_{\epsilon_h}^2 = 1$, which are within the typical ranges of parameter values observed in the motivating example as well as other empirical studies in psychology utilizing variations of the VAR model (Chow et al., 2007,0). The two covariates, x_{1nt} and x_{2nt} , were generated according to the following models:

$$\begin{aligned} p_{nt} &= \text{logit}^{-1}(\beta_1 + \beta_2 v_{1nt}) \\ x_{1nt} &\sim \text{Bern}(p_{nt}) \\ x_{2nt} &= \psi_1 + \psi_2 x_{1nt} + \psi_3 v_{1nt} + \epsilon_{x_2} \\ \epsilon_{x_2} &\sim \mathcal{N}(0, \sigma_{\epsilon_{x_2}}^2), \end{aligned} \tag{11}$$

As we can see from Equations 11, the first covariate, denoted x_1 , was a binary variable that followed a Bernoulli distribution at each observational time point, with probability of observing values of 1 dependent on contemporaneous observations being p_{nt} , which is parameterized by $\beta_1 = 0$, $\beta_2 = -2$. In Equations 11, v_1 represents a fully observed auxiliary variable with a uniform distribution over $(-3, 3)$. The second covariate, denoted x_2 , followed a normal distribution at each time point with means derived from contemporaneous observations of both v_1 and x_1 , and parametrized via $\psi_1 = 2$, $\psi_2 = -0.2$, $\psi_3 = 0.2$, and $\sigma_{\epsilon_{x_2}}^2 = 0.25$.

Missingness Settings

We simulated missingness in both the dependent variables and the covariates following each of the three possible missing data mechanisms: MCAR, MAR and NMAR. Under NMAR, four combinations of NMAR strength and missing value percentage were considered. Let $r_{wnt}, r_{hnt}, r_{x_{1nt}}, r_{x_{2nt}}$ be the missingness indicators for the dependent

variables and the covariates, respectively, such that $r_{nt}=1$ if the corresponding variable for dyad n at time t is missing and 0 otherwise. Hence, r_{nt} was Bernoulli distributed with probability of missing $P(r_{nt})$, which was determined by different variables depending on the nature of the missingness. The most general missing data model considered was an NMAR model in which we specified the probability of the r_{nt} as conditional on the variables subjected to missingness themselves (h_{nt} , w_{nt} , x_{1nt} , or x_{2nt}) and fully observed variables v_{2nt} and v_{3nt} .

The missing data model for the dependent variables is expressed as:

$$\begin{aligned} P(r_{w_{nt}} = 1 | w_{nt}, v_{2nt}, v_{3nt}) &= \text{logit}^{-1}(\phi_{inty_1} + \phi_y w_{nt} + \phi_{v_2} v_{2nt} + \phi_{v_3} v_{3nt}) \\ P(r_{h_{nt}} = 1 | h_{nt}, v_{2nt}, v_{3nt}) &= \text{logit}^{-1}(\phi_{inty_2} + \phi_y h_{nt} + \phi_{v_2} v_{2nt} + \phi_{v_3} v_{3nt}), \end{aligned} \quad (12)$$

where v_2 and v_3 were both uniformly distributed over $(0, 3)$.

We also generated missingness for the covariates according to the functions below:

$$\begin{aligned} P(r_{x_{1nt}} = 1 | x_{1nt}, v_{2nt}, v_{3nt}) &= \text{logit}^{-1}(\phi_{intx_1} + \phi_{x_1} x_{1nt} + \phi_{v_2} v_{2nt} + \phi_{v_3} v_{3nt}) \\ P(r_{x_{2nt}} = 1 | x_{2nt}, v_{2nt}, v_{3nt}) &= \text{logit}^{-1}(\phi_{intx_2} + \phi_{x_2} x_{2nt} + \phi_{v_2} v_{2nt} + \phi_{v_3} v_{3nt}), \end{aligned} \quad (13)$$

In the scenario of our empirical example, this would capture the possibility that one's perception of conflict resolution affects his/her probability of reporting of such conflict, and also that the reporting of a conflict is affected by other factors such as the emotional state one is in (fully observed variables). In terms of covariates, children's reactions can be missing due to a lack of reaction (NMAR), or an absence of parent's reporting (MAR). We made two assumptions regarding the nature of the missing data mechanisms. First, we assumed that the effect of w_{nt} on $P(r_{w_{nt}})$ and the effect of h_{nt} on $P(r_{h_{nt}})$ were the same (as represented by a sole ϕ_y in both equations in 12) given that they were measured the same way and of the same construct, just for different individuals. Second, we assumed that the effects of fully observed variables v_{2nt} and v_{3nt} were the same for all variables (as represented by only one set of ϕ_{v_2} and ϕ_{v_3} in Equations 12 and 13). In other words, each unit of increase in these fully observed variables was assumed to affect all variables' log

odds of being missing on a particular occasion by the same degree. The missing data model was constructed in this way to strike a balance between offering a reasonable missing data scenario in the study of affects, and realistic levels of model complexity from an estimation standpoint. In addition, it mirrors the missing data patterns observed in many ecological momentary assessment studies in that individuals tend to show heightened probability of omitting responses to an entire set of items (e.g., sections of a survey, or even an entire survey), as opposed to isolated items on the survey.

With the aforementioned general NMAR missing data model, the MAR and MCAR models in this simulation can be viewed as reduced special cases of the general model. To simulate missingness according to the MAR mechanism, we set ϕ_y , ϕ_{x_1} and ϕ_{x_2} in Equations 12 and 13 to 0 so that missingness did not depend on the dependent variables in focus, but on the two fully observed variables, v_2 and v_3 . For the MCAR condition, all ϕ 's except for the intercept term were set to be 0 in the model, thus the missingness did not depend on any data-related information. Table 3 contains all the parameter values used in missingness generation. For MCAR and MAR missingness, we simulated roughly 50% of missingness in each variable. For the NMAR, we investigated four different conditions: 1. where impact of NMAR factors (w_{nt} , h_{nt} , x_{1nt} , and x_{2nt}) were relative low and each variable had roughly 30% of data missing ("NMARL3"), 2. where impact of NMAR factors were relative low and each variable had roughly 50% of data missing ("NMARL5"), 3. where impact of NMAR factors were relative high and each variable had roughly 30% of data missing ("NMARH3"), 4. where impact of NMAR factors were relative high and each variable had roughly 50% of data missing ("NMARH5"). The ϕ values were chosen to ensure the missingness percentages and NMAR impact in each condition.

Model Fitting Procedures

Each of the datasets simulated under different missing data mechanisms and percentages were fitted to a bivariate VAR model using different missing data handling approaches, including the LD method, two-step partial MI method, and Bayesian

approach. With the LD method, observations with any missing variables were excluded from the analysis. The listwise deleted data were then fitted to the VAR model using a Bayesian approach. This method was considered as the baseline method. For the two-step partial MI method, missing data in the covariates were multiply imputed, and model fitting procedures were performed using frequentist method, specifically, by optimizing a log-likelihood function constructed using prediction errors obtained from running the Kalman filter (Chow et al., 2010b; Schweppe, 1965).

Under the Bayesian approach, Bayesian models with MCAR, MAR, and NMAR missing data model components were fitted to all data sets, regardless of true missing data generation mechanisms, resulting in 18 conditions for the Bayesian missing data approach (6 true data conditions \times 3 Bayesian models for model fitting). In the remaining sections, we refer to results from the Bayesian approach by *true data generation conditions \times fitted model*. For instance, NMARL5 \times MAR stands for low dependency NMAR data with 50% of missingness fitted to a Bayesian model with MAR missing data model, specifically the missing data model in Equation 12 - 13 with the NMAR parameters, ϕ_y , ϕ_{x_1} , and ϕ_{x_2} , fixed at 0 as opposed to freely estimated.

We categorized fitted models in the Bayesian approach into three scenarios: correctly specified models, underfitting models (when the model only contains part of the true model), and overfitting models (when the true model can be seen as a special case of the fitted model with some parameters having values of 0). Correctly specified models included MCAR \times MCAR, MAR \times MAR, and NMAR \times NMAR. Underfitting models were MAR \times MCAR, NMAR \times MCAR, and NMAR \times MAR. MCAR \times MAR, MCAR \times NMAR, and MAR \times NMAR are all considered as overfitting models. Response variables of interest and covariates that contained missingness were modeled with the true data generating models (Equations 2 and 11).

Consistent with the approach adopted in the empirical illustration, weakly informative priors were set for the parameters in the VAR model. For the rest of the parameters, the

prior settings were non-informative. Given the complexity of the complete data models, we used the following setting in sampling: 2 chains, 1000 steps of adaptation, 5000 steps of burn-in, and 200000 samples per chain.

The partial MI method was adopted for the two-step MI approach in this paper. Specifically, we first generated $m=5$ sets of plausible values of missing covariates. Then, the missing observations of the covariates in the simulated data sets were filled in with the imputed values, resulting in five data sets with fully observed covariates and missing data in dependent variables. Those data sets were used in VAR model fitting, and the estimated parameters from the five imputed data sets were pooled together based on Rubin's method (Rubin, 1996). More details about the partial MI was described in Ji et al. (2018). With the partial MI method, an inclusive model, which includes all available information from the data set, is usually recommended for imputation procedures. Therefore, regardless of missing data conditions, the same imputation model was applied, making use of all variables involved in the data/missingness generation procedures, and lagged variables as necessary. To be specific, the imputation model included fully observed variables v_{1nt} , v_{2nt} , and v_{3nt} , model covariates with missing values x_{1nt} and x_{2nt} , and dependent variables and their values at the immediate previous time points w_{nt} , h_{nt} , $w_{n,t-1}$, and $h_{n,t-1}$. VAR model fitting under the partial MI approach was performed using the dynr package in R (Ouet al., 2017), which provides an option for fitting discrete-time state-space models such as the VAR model by optimizing a log-likelihood function constructed using prediction errors obtained from running the Kalman filter on a specified state-space model (Chow et al., 2010b; Schweppe, 1965).

Performance Measures

In order to compare between a Bayesian approach and a frequentist approach, we treated the Bayesian posterior mean of a parameter θ as the point estimate for that parameter ($\hat{\theta}$), and the posterior standard deviation as the standard error of that estimate. With that, we compared the following performance measures for each parameter: 1.

relative bias (Equation 14), 2. root-mean-square error (RMSE, Equation 15), 3. difference in average standard error across MC runs compared to the MC standard deviation (MCSDs³; “dSD”, Equation 16), 4. difference in average standard error across MC runs compared to the MCSD from the fitting the model for dependent variables (Equation 2 only) to the complete data without missingness (“dSDFull”), 5. coverage, defined as percentage of MC runs in which the credible or confidence intervals contained the true parameter value. Relative bias and RMSE, in particular, are defined as:

$$\text{relative bias}(\theta) = \frac{1}{H} \sum_{h=1}^H \left(\frac{\hat{\theta}_h - \theta}{\theta} \right) \quad (14)$$

$$\text{RMSE}(\theta) = \sqrt{\frac{1}{H} \sum_{h=1}^H (\hat{\theta}_h - \theta)^2}, \quad (15)$$

$$\text{dSD}(\theta) = \frac{1}{H} \sum_{h=1}^H \text{SE}_h(\hat{\theta}) - \sqrt{\frac{1}{H-1} \sum_{h=1}^H (\hat{\theta}_h - \bar{\theta})^2}, \quad (16)$$

where H represents the total number of MC runs, and $\bar{\theta} = \sum_{h=1}^H \hat{\theta}_h / H$. Finally, we also evaluated the performance of the BIC, as defined in Equation 9, as one possible model selection measure.

Simulation Results

To facilitate comparisons among parameters that appeared in different submodels, we grouped VAR parameters into three categories: dynamic ($a_{w \rightarrow w}$, $a_{h \rightarrow h}$, $b_{h \rightarrow w}$, $b_{w \rightarrow h}$), covariate ($c_{x_1 \rightarrow w}$, $c_{x_1 \rightarrow h}$, $d_{x_2 \rightarrow w}$, $d_{x_2 \rightarrow h}$) and noise ($\sigma_{\epsilon_w}^2$, $\sigma_{\epsilon_h}^2$) parameters, similar to the grouping done in Ji et al. (2018). Figure 2 shows the average relative bias and RMSE of estimates within each group across conditions. In general, as expected, both the two-step partial MI approach and the Bayesian selection model approach outperformed the LD approach, which generated biased results for all parameters under all missing data conditions. The advantage of the Bayesian approach was more apparent under the NMAR conditions, especially with higher percentage of missingness and higher NMAR dependency. Compared with the percentage of missingness, NMAR dependency was

³ The standard deviation of all point estimates on a parameter across MC runs

observed to have more notable effects on parameter estimation accuracy. We now elaborate on specific details pertaining to each missing data handling technique.

Simulation results for the LD approach. The LD approach generated biased parameter point estimates, as evaluated by relative bias and RMSE under all conditions. Compared to covariate parameters and noise parameters, the dynamic parameters were affected the most when the LD approach was used because deleting missing observations led to distortion in the time intervals between successive measurements. Figure 2 shows that dynamic parameters tended to be under-estimated with LD approach, with negative relative biases, under all conditions. Covariate parameters were not as affected by LD as the other two parameter groups in terms of relative biases in point estimates, probably due to the fact that relationships between covariates and the variables of interest would still be preserved under LD. An interesting observation was that the RMSEs for covariates were visibly larger with LD across all conditions with 50% missing data (i.e., MCAR, MAR, NMARL5, NMARH5) compared to when other methods were used, suggesting that estimation biases in covariates using LD were due more to the influence of missingness percentages, regardless of missingness mechanisms. Such biases in parameter point estimates also resulted in low coverage rate (i.e., around 50% when averaged across all VAR parameters) for the LD approach under all conditions.

In terms of standard error estimates, using LD approach in general yielded very similar results when compared with empirical MC standard deviation, reflected by close to zero dSD across all conditions. However, when compared with model estimation results using full data set without missingness, LD approach led to over-estimation of standard errors for all parameters under all conditions. Since LD approach includes fewer observations in the analysis, inflation of standard error estimates is in accordance with our expectation.

A comparison of the results using two-step partial MI and Bayesian approaches. Under MCAR and MAR conditions, using approaches other than LD yielded comparable biases in estimates. For dynamic and covariate parameters, the two-step

partial MI approach and the Bayesian approach both yielded very small biases and RMSEs, regardless of whether the missingness model was correctly specified, underfitting or overfitting with the Bayesian approach. Figure 2 also shows that estimates for noise parameters had higher biases and RMSEs in the two-step partial MI approach compared to those from the three modeling conditions in the Bayesian approach, with results in the Bayesian approach being very similar to each other. This may be due to the fact that in the two-step partial MI approach, missing data imputation models were slightly misspecified.

Table 4 contains comparisons of dSDs and dSDFulls across conditions in the three parameter groups. Except for those resulted from LD, all the dSDs and dSDFulls were small and similar to each other, suggesting that the standard error estimates from the two-step approach and the posterior standard deviations from the Bayesian approach both approximated the degree of sampling variability in the parameters well under MCAR and MAR conditions. The coverage as averaged across parameters were comparable for the two-step partial MI and Bayesian approaches, and very similar for the MCAR and MAR conditions. To be specific, the coverage for the Bayesian approach were slightly better, with better point estimation results for noise parameters, resulting in an average coverage of 95% across all parameters for both the MCAR and MAR conditions. The average coverage across all parameters for the two-step partial MI approach was 86% for the MCAR condition and 87% for the MAR condition.

Under NMAR conditions, clearer differences emerged between the two-step partial MI and the Bayesian approach. When missingness model was correctly specified (i.e., NMAR×NMAR) using the Bayesian method, dynamic parameters were well recovered with close to zero relative bias and very small RMSE, even when the NMAR dependency was high and 50% of observations were missing (i.e., in the NMARH5 condition). The two-step partial MI approach resulted in slightly higher biases and RMSEs compared to the Bayesian approach when a correctly specified missing data model was fitted, especially with increase in NMAR strength and missing data percentage. In contrast, when the

missing data model was an underfitting one (e.g., a MCAR or MAR model fitted to NMAR data), the Bayesian approach yielded slightly higher biases and RMSEs than the two-step partial MI approach, which used the same imputation model that allowed for NMAR elements regardless of the true missing data mechanism. The differences between the Bayesian and two-step partial MI approaches grew as the NMAR dependency and the percentage of missingness increased.

Compared to covariate parameters and process noise parameters, dynamic parameters were the least affected by misspecification of missing data model in that underfitting missingness models (i.e., NMAR \times MCAR, or NMAR \times MAR) only resulted in a very small addition in biases and RMSEs compared to the true model (i.e., NMAR \times NMAR) across all NMAR conditions. Simulation results suggested that the differences in biases were more affected by NMAR dependency (e.g. comparing NMARL3 and NMARH3) than the percentage of missingness (e.g. comparing NMARL3 and NMARL5).

The aforementioned simulation results regarding parameter point estimates were also reflected in the credible intervals resulted from the Bayesian approaches and confidence intervals from the two-step partial MI approach. We plotted the coverage results only from the NMAR conditions because the differences between missing data handling techniques were the most salient under these conditions (Figure 3). The upper graph shows the overall coverages across all VAR parameters, and the lower graph grouped coverage percentages by parameter type. Under NMAR, only the Bayesian approach with correctly specified missing data model resulted in coverage percentages that were close to the nominal 95% across all parameters. With the exception of the LD method, coverage of the dynamic parameters was largely similar across different missing data handling methods and was not very sensitive to misspecification in the missing data model. However, the coverage percentages for covariate and noise parameters were considerably worse when using LD, the two-step partial MI, or with an underfitting model in the Bayesian approach. The coverage percentages continued to worsen with increase in NMAR dependency and missing data

percentage.

To summarize, the LD approach produced the most biased point estimates, especially for dynamic parameters under all missing data conditions. Standard error estimates were also larger with the LD approach, as compared with standard errors when full data was fitted to the model. The Bayesian approach, when coupled with the correctly specified missing data model, performed the best under NMAR conditions. Even with 50% of missing data and relatively high NMAR dependency, the parameters were well recovered and the coverage rates were close to the 95% nominal level. The two-step partial MI approach also produced reasonable estimation results under MCAR and MAR conditions. Under NMAR conditions, the two-step partial MI approach generated similar, and even slightly better results than the Bayesian approach when underfitting models were used (i.e., NMAR \times MCAR, NMAR \times MAR), as indicated by slightly smaller RMSE and better coverage rate for the two-step partial MI approach results, when compared with underfitting Bayesian approaches. In addition, the biases in estimation for the two-step partial MI approach and underfitting Bayesian models were more substantial for covariate and noise parameters, than for dynamic parameters.

Missingness Model Parameters and Model Selections. Table 5 compares parameter estimates from fitting the dynamic model with a MCAR, MAR and NMAR missingness model to the true parameter values under the six missingness scenarios (MCAR, MAR, NMARL3, NMARL5, NMARH3, NMARH5, respectively) ⁴. As expected, fitting the correctly specified missingness model yielded good accuracy in the estimated missing data parameter values. A misspecified but overfitting missingness model (i.e., MAR \times NMAR, MCAR \times NMAR, or MCAR \times MAR) also yield satisfactory estimation results, with the overfitted parameters estimated to be close to 0 (i.e., credible intervals centered around 0, as opposed to the ones for other nonzero missingness parameters where

⁴ Since neither the LD method nor the two-step partial MI method explicitly specify and estimate a missingness model, no missing data parameter estimates were available from these methods.

credible intervals did not contain 0). Looking at credible intervals therefore allows us to infer the missingness mechanism, especially in terms of NMAR missingness versus MAR or MCAR. Underfitting regarding the missingness mechanism, however, resulted in biased estimates in the missing data model parameters. For instance, when an MAR model was fitted to NMAR data, estimates of the intercept as well as the MAR parameters all showed notable biases.

A comparison of the BIC values under all true missingness mechanisms and missingness models can be found in Figure 4. The BIC performed well in selecting the true missingness model (i.e., BIC was the lowest for the true data generating model among all missingness models 100% out of all the trials). Model selection results based on the BIC were consistent with the conclusions drawn from inspecting the credible intervals of the missing data parameters. For instance, when the true missing data mechanism was NMAR, the BIC suggested — across all the trials considered — a preference for the NMAR model over other missing data models. Consistent with results from the BIC, the corresponding credible intervals for NMAR parameters, namely ϕ_{x_1} , ϕ_{x_2} , and ϕ_y , also indicated that these parameters deviated substantially from zero.

Discussion

Missing data, especially NMAR missingness and missing observations in covariates, pose great challenges in model estimation for ILD. We applied a selection model approach with Bayesian estimation procedures to fit both the model of interest, a bivariate VAR model with covariates, and the missing data model simultaneously to time series data with missingness. We performed a simulation study to illustrate the different performances of the Bayesian approach and MI approach under different missing data mechanisms, including MCAR, MAR, and NMAR. In order to demonstrate how percentages of missing data, and NMAR dependencies may affect the estimation results, we tested four NMAR conditions, including low dependency NMAR and high dependency NMAR paired with 30% and 50% of missingness. Simulation results showed that the performance of the Bayesian

approach was satisfactory across all tested conditions when correct missing data model was fitted, with close to zero relative biases and very small RMSEs (i.e., smaller than .06). The standard deviation estimates were also very close to empirical standard deviations across 100 MC runs, as well as the standard deviations obtained when the complete data were used for model fitting (i.e., differences smaller than .03 for all parameters). The coverage for all parameters were very close to 95%. In addition, with correctly specified missingness model, all missingness model parameters were reliably recovered. The application of the Bayesian modeling approach was illustrated with an empirical data set and the estimation results were comparable with previous studies. Compared with the Bayesian approach with correct missing data model, the two-step partial MI approach generated relatively reliable estimates for dynamic parameters across all conditions, but parameter point estimates for covariate and process noise parameters were biased under NMAR conditions, resulting in relatively low coverage. This finding was consistent with the finding in the [Ji et al. \(2018\)](#) paper. When underfitting Bayesian models (i.e., NMAR \times MCAR, NMAR \times MAR) were used with NMAR conditions, the two-step partial MI approach generated similar, and even slightly better results than the Bayesian approach.

When the Bayesian selection model approach is used to model missingness mechanism, implementing an overfitting missing data model can have some advantages in some contexts. For instance, if the model is not too highly parameterized, the overfitting model can still lead to correct inferences regarding other important parameters in the model. Doing so also allows researchers to examine the plausible missing data mechanism and infer whether the missingness is MCAR, MAR or NMAR. It is also possible to obtain posterior distributions on the missing data, which may be of substantive interest.

It is also worth pointing out that as a parametric likelihood method, this approach requires good knowledge of both the model for the dependent variables and the missingness model. Misspecified models or overly complicated models can easily lead to convergence issue. For instance, when applying the Bayesian selection model approach to the empirical

data in this paper, we took several approaches to simplify the missing data model in order to achieve model convergence, which included constraining most of the missing data parameter values to be the same for the covariate parameters. Given the theoretically similar missing data mechanisms for the two covariates in the model, we considered such constraints appropriate. However, in other empirical studies, different missing data models may be necessary for different variables with missingness. In addition, the fact that in real life, we never know what is the true missing data model may also pose challenge in model fitting of the Bayesian selection model approach. Furthermore, with complex models and large data sets, the Bayesian method may be more time-consuming compared with fully or partially frequentist methods, such as the two-step partial MI approach. With relatively moderate percentage of missingness (i.e., 30%) and low NMAR dependency level, the two-step partial MI approach were able to recover most of the parameters of the model of interest, and the model fitting procedures take relatively short period of time (115.81 seconds for MI vs. 47697.73 seconds in Bayesian given the specifications of iterations provided).

Similar to other dynamic models for longitudinal data, to start the process, initial conditions need to be first specified. The specification we adopted in this article coincides with what is typically regarded as “diffused initial condition” in the state space model literature. A diffused initial condition typically involves specifying the distributions of the latent variables at the first time point to be approximated by parametric distributions that are “spread-out” enough to include the true initial conditions, such as a series of univariate normal distributions with initial means at zeros and arbitrarily large variances (e.g., [De Jong, 1991](#); [Harvey, 2001](#); [Schweppe, 1973](#)). This was the initial condition distribution adopted in this article, and the variance we imposed (100) was chosen to ensure that the majority of the posterior probability distribution is within the possible range of values for the dependent variables w and h . Other approaches exist in addressing the first time points in dynamic modeling, including freely estimating the mean and covariance structure (e.g.

Harvey and Souza, 1987; McArdle and Epstein, 1987), or using the model-implied mean and covariance (e.g. du Toit and Browne, 2007; Hamaker, 2005), among a few. The specification of initial time points is itself an active area of research that is beyond the scope of this article (e.g. du Toit and Browne, 2007; Ji et al., 2018; Ou et al., 2016).

Our simulation results also confirmed the utility of the BIC in selecting the correct missing data model. Model selection under the Bayesian framework with information criteria is still an active field of research. The utilization of information criteria should not be restricted to BIC, and the way we computed BIC is only one way of deriving it from posterior samples. BIC was computed in the aforementioned way primarily for computational ease, in that it did not require additional MCMC sampling or explicit evaluation of the log-likelihood. Future research should explore a more diverse pool of information criteria, some examples being the conditional Deviance Information Criterion (DIC) considered by Celeux et al. (2006), and the modified DIC, which is a combination of the BIC and the DIC, proposed by Lu and Song (2012).

This article demonstrated a Bayesian selection model approach to handle missing data in longitudinal studies when a bivariate VAR model with covariates, a very widely applied model in empirical studies, was fitted. In our study, the covariance of the process noises in the VAR model was fixed at zero and thus not estimated. In other applications, it may be of interest to free up the covariance by using a bivariate normal prior distribution specification in the Bayesian method. In addition, further simulation studies need to be conducted to evaluate the performance of this Bayesian selection model approach with different full-data model specification. For instance, with Bayesian estimation approach, a natural extension would be longitudinal models with multilevel structure with random effects. Since MI methods have also been developed to handle nested data, it would be interesting to compare the performance of the Bayesian approach and multilevel MI approach. It would also be of interest to evaluate and compare the performance of the Bayesian selection model approach and MI approaches with other dynamic model, such as

VAR models with categorical data, process factor analysis models, regime switching models.

Table 1

Empirical Estimation Results Using the Bayesian Approach with an NMAR Missingness Model and Using the partial MI approach

Parameter	Bayesian		MI	
	Mean	SD	Mean	SE
VAR model parameters				
$a_{w \rightarrow w}$	0.0037	0.0377	-0.028	0.047
$a_{h \rightarrow h}$	-0.0417	0.0382	0.017	0.048
$b_{h \rightarrow w}$	0.0171	0.0377	0.039	0.046
$b_{w \rightarrow h}$	0.0585	0.0382	-0.054	0.047
$c_{x_1 \rightarrow w}$	-0.5035	0.0178	-0.121	0.024
$c_{x_1 \rightarrow h}$	-0.5096	0.0178	-0.132	0.026
$d_{x_2 \rightarrow w}$	0.0623	0.0184	0.033	0.041
$d_{x_2 \rightarrow h}$	0.0608	0.0186	0.017	0.035
$\sigma_{\epsilon_w}^2$	0.5449	0.0207	0.928	0.032
$\sigma_{\epsilon_h}^2$	0.5279	0.0203	0.917	0.032
Covariate model parameters				
β_1	0.4233	0.1218		
β_2	-0.2937	0.0150		
ψ_1	2.1743	0.4363		
ψ_2	0.8212	0.1873		
ψ_3	0.2402	0.0625		
ψ_4	0.0674	0.0364		
$\sigma_{\epsilon_{x_1}}^2$	1.284	0.1056		
$\sigma_{\epsilon_{x_2}}^2$	7.0253	0.6747		
Missingness model parameters				
$\phi_{int_{y_1}}$	-5.9096	0.4406		
$\phi_{int_{y_2}}$	-5.7389	0.4093		
$\phi_{int_{x_1}}$	-10.5266	1.2860		
$\phi_{int_{x_2}}$	-10.5264	1.2853		
$\phi_{p \rightarrow y}$	0.181	0.1208		
$\phi_{p \rightarrow x}$	0.0954	0.1067		
ϕ_x	5.3734	0.5267		

Note: A number in bold indicates that the credible/confidence interval for that estimate did not contain zero.

Table 2
Sensitivity Check Results for Empirical Illustration

Parameter	VAR1	VAR2	VAR3	Cov1	Cov2	Cov3	NA1	NA2	NA3
VAR model parameters									
$a_{w \rightarrow w}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$a_{h \rightarrow h}$	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04
$b_{h \rightarrow w}$	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
$b_{w \rightarrow h}$	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
$c_{x_1 \rightarrow w}$	-0.50	-0.50	-0.50	-0.50	-0.50	-0.50	-0.50	-0.50	-0.50
$c_{x_1 \rightarrow h}$	-0.51	-0.51	-0.51	-0.51	-0.51	-0.51	-0.51	-0.51	-0.51
$d_{x_2 \rightarrow w}$	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
$d_{x_2 \rightarrow h}$	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
$\sigma_{\epsilon_w}^2$	0.54	0.55	0.54	0.54	0.54	0.55	0.54	0.54	0.55
$\sigma_{\epsilon_h}^2$	0.53	0.53	0.52	0.53	0.53	0.53	0.53	0.53	0.53
Covariate model parameters									
β_1	0.43	0.42	0.43	0.42	0.43	0.41	0.43	0.43	0.46
β_2	-0.29	-0.29	-0.29	-0.29	-0.29	-0.29	-0.29	-0.29	-0.29
ψ_1	2.16	2.17	2.17	2.18	2.03	2.51	2.16	2.15	2.25
ψ_2	0.83	0.82	0.83	0.82	0.83	0.82	0.83	0.83	0.91
ψ_3	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.27
ψ_4	0.07	0.07	0.07	0.07	0.08	0.04	0.07	0.07	0.07
$\sigma_{\epsilon_{x_1}}^2$	1.29	1.28	1.30	1.28	1.29	1.27	1.29	1.29	1.31
$\sigma_{\epsilon_{x_2}}^2$	6.98	7.05	6.98	7.03	6.96	7.15	7.02	6.86	7.65
Missingness model parameters									
$\phi_{int_{y_1}}$	-5.91	-5.91	-5.91	-5.91	-5.91	-5.91	-5.92	-5.86	-6.05
$\phi_{int_{y_2}}$	-5.74	-5.74	-5.74	-5.74	-5.74	-5.74	-5.75	-5.70	-5.87
$\phi_{int_{x_1}}$	-10.49	-10.54	-10.47	-10.54	-10.59	-10.67	-10.72	-9.54	-16.41
$\phi_{int_{x_2}}$	-10.49	-10.54	-10.47	-10.54	-10.58	-10.67	-10.72	-9.54	-16.41
$\phi_{p \rightarrow y}$	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.20
$\phi_{p \rightarrow x}$	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.09	0.10
ϕ_x	5.36	5.38	5.35	5.38	5.40	5.42	5.45	4.98	7.63

Notes: VAR: VAR model parameters check; Cov: Covariate model parameters check; NA: Missingness model parameters check.

Table 3

Parameter Values for Missing Data Mechanisms in Simulation Study

	MCAR	MAR	NMARL3	NMARL5	NMARH3	NMARH5
ϕ_{intx_1}	0	-1.05	-3.04	-2.19	-4.18	-3.33
ϕ_{intx_2}	0	-1.05	-2.40	-1.55	-2.90	-1.45
ϕ_{inty_1}	0	-1.05	-2.45	-1.60	-3.05	-2.20
ϕ_{inty_2}	0	-1.05	-2.35	-1.50	-2.81	-1.96
ϕ_{v_2}	0	0.30	0.30	0.30	0.30	0.30
ϕ_{v_3}	0	0.40	0.40	0.40	0.40	0.40
ϕ_{x_1}	0	0	0.60	0.60	1.20	1.20
ϕ_{x_2}	0	0	1.00	1.00	2.00	2.00
ϕ_y	0	0	0.50	0.50	1.00	1.00

Table 4
Standard Deviation/Error Estimates Comparison for Simulation Study

Method used/ Model fitted	True: MCAR						True: MAR					
	dSD			dSDFull			dSD			dSDFull		
	Dynamic	Covariate	Noise	Dynamic	Covariate	Noise	Dynamic	Covariate	Noise	Dynamic	Covariate	Noise
LD	0.00	0.00	0.02	0.06	0.10	0.13	-0.00	-0.01	0.01	0.05	0.08	0.11
MI	0.00	-0.00	0.00	0.01	0.02	0.02	0.00	0.00	-0.00	0.01	0.02	0.02
MCAR	-0.00	-0.00	0.00	0.01	0.02	0.01	0.00	0.00	-0.00	0.01	0.02	0.01
MAR	-0.00	-0.00	0.00	0.01	0.02	0.01	0.00	0.00	-0.00	0.01	0.02	0.01
NMAR	-0.00	-0.00	0.00	0.01	0.02	0.01	0.00	0.00	-0.00	0.01	0.02	0.02

Method used/ Model fitted	True: NMARL3						True: NMARL5					
	dSD			dSDFull			dSD			dSDFull		
	Dynamic	Covariate	Noise	Dynamic	Covariate	Noise	Dynamic	Covariate	Noise	Dynamic	Covariate	Noise
LD	-0.00	0.00	-0.01	0.02	0.03	0.04	0.00	0.00	-0.00	0.06	0.09	0.10
MI	-0.00	0.00	-0.00	0.01	0.01	0.01	0.00	0.00	-0.00	0.02	0.02	0.02
MCAR	-0.00	0.00	-0.00	0.01	0.01	0.01	-0.00	0.00	-0.00	0.01	0.02	0.01
MAR	-0.00	0.00	-0.00	0.01	0.01	0.01	-0.00	0.00	-0.00	0.01	0.02	0.01
NMAR	-0.00	0.00	-0.00	0.01	0.01	0.01	-0.00	0.00	-0.00	0.01	0.02	0.02

Method used/ Model fitted	True: NMARH3						True: NMARH5					
	dSD			dSDFull			dSD			dSDFull		
	Dynamic	Covariate	Noise	Dynamic	Covariate	Noise	Dynamic	Covariate	Noise	Dynamic	Covariate	Noise
LD	-0.00	0.00	-0.00	0.03	0.04	0.04	-0.00	0.01	0.00	0.07	0.11	0.11
MI	0.00	0.00	-0.00	0.01	0.01	0.00	0.00	0.00	-0.00	0.02	0.02	0.01
MCAR	0.00	0.00	-0.00	0.01	0.01	0.00	0.00	0.00	-0.00	0.02	0.01	0.01
MAR	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	-0.00	0.02	0.01	0.01
NMAR	0.00	0.00	0.00	0.01	0.01	0.02	0.00	0.00	-0.00	0.01	0.02	0.03

Note: dSD: difference in average standard error across MC runs compared MCSDs; dSDFull: difference in average standard error across MC runs compared to the MCSD from the fitting the model for dependent variables to the complete data without missingness.

Table 5
Estimated Missingness Parameter Values for Simulation Study

Parameter	MCAR				MAR				NMARL3			
	True	MCAR	MAR	NMAR	True	MCAR	MAR	NMAR	True	MCAR	MAR	NMAR
ϕ_{intx_1}	0	-0.01	-0.01	-0.01	-1.05	0.01	-1.04	-1.06	-3.04	-0.78	-1.82	-3.08
ϕ_{intx_2}	0	-0.01	-0.01	0	-1.05	0	-1.04	-1.04	-2.4	-0.78	-1.81	-2.41
ϕ_{inty_1}	0	0	0	-0.01	-1.05	0	-1.05	-1.05	-2.45	-0.58	-1.61	-2.47
ϕ_{inty_2}	0	0.01	0.01	0	-1.05	0	-1.04	-1.05	-2.35	-0.78	-1.81	-2.37
ϕ_{v_2}			0	0	0.3		0.3	0.3	0.3		0.29	0.3
ϕ_{v_3}			0	0	0.4		0.4	0.4	0.4		0.38	0.4
ϕ_{x_1}				0				0.01	0.6			0.62
ϕ_{x_2}				-0.02				-0.01	1			1
ϕ_y				0				0	0.5			0.51
Parameter	NMARL5				NMARH3				NMARH5			
	True	MCAR	MAR	NMAR	True	MCAR	MAR	NMAR	True	MCAR	MAR	NMAR
ϕ_{intx_1}	-2.19	0	-0.98	-2.22	-4.18	-0.73	-1.62	-4.17	-3.33	0	-0.86	-3.36
ϕ_{intx_2}	-1.55	0	-0.98	-1.54	-2.9	-0.66	-1.56	-2.91	-1.45	0.47	-0.38	-1.46
ϕ_{inty_1}	-1.6	0.17	-0.81	-1.6	-3.05	-0.4	-1.29	-3.05	-2.2	0.24	-0.61	-2.24
ϕ_{inty_2}	-1.5	0	-0.99	-1.51	-2.81	-0.69	-1.58	-2.81	-1.96	-0.01	-0.87	-1.98
ϕ_{v_2}	0.3		0.28	0.3	0.3		0.25	0.3	0.3		0.25	0.3
ϕ_{v_3}	0.4		0.38	0.4	0.4		0.33	0.4	0.4		0.33	0.4
ϕ_{x_1}	0.6			0.61	1.2			1.19	1.2			1.21
ϕ_{x_2}	1			0.98	2			2.01	2			2.02
ϕ_y	0.5			0.51	1			0.99	1			1.01

Note: A number in bold indicates that over 95% of credible intervals for that estimate did not contain zero within a specific simulation condition.

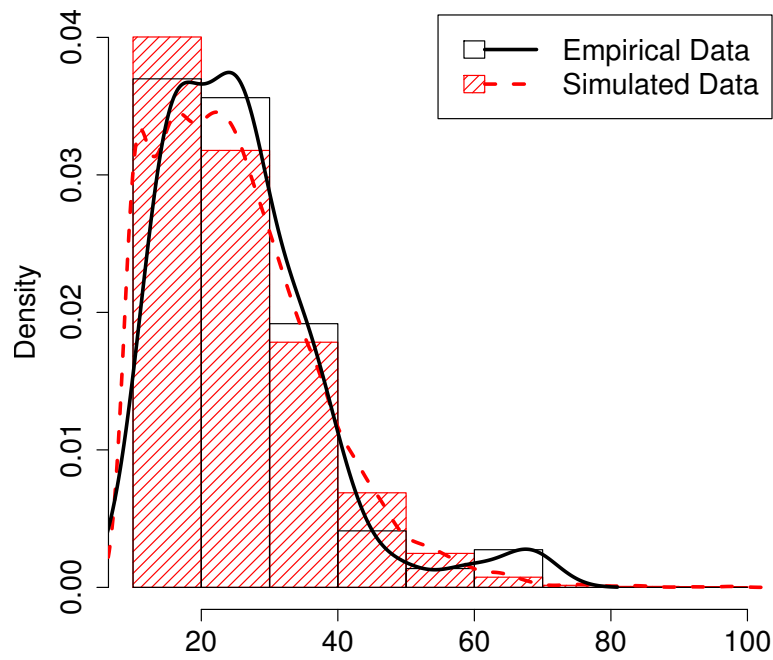
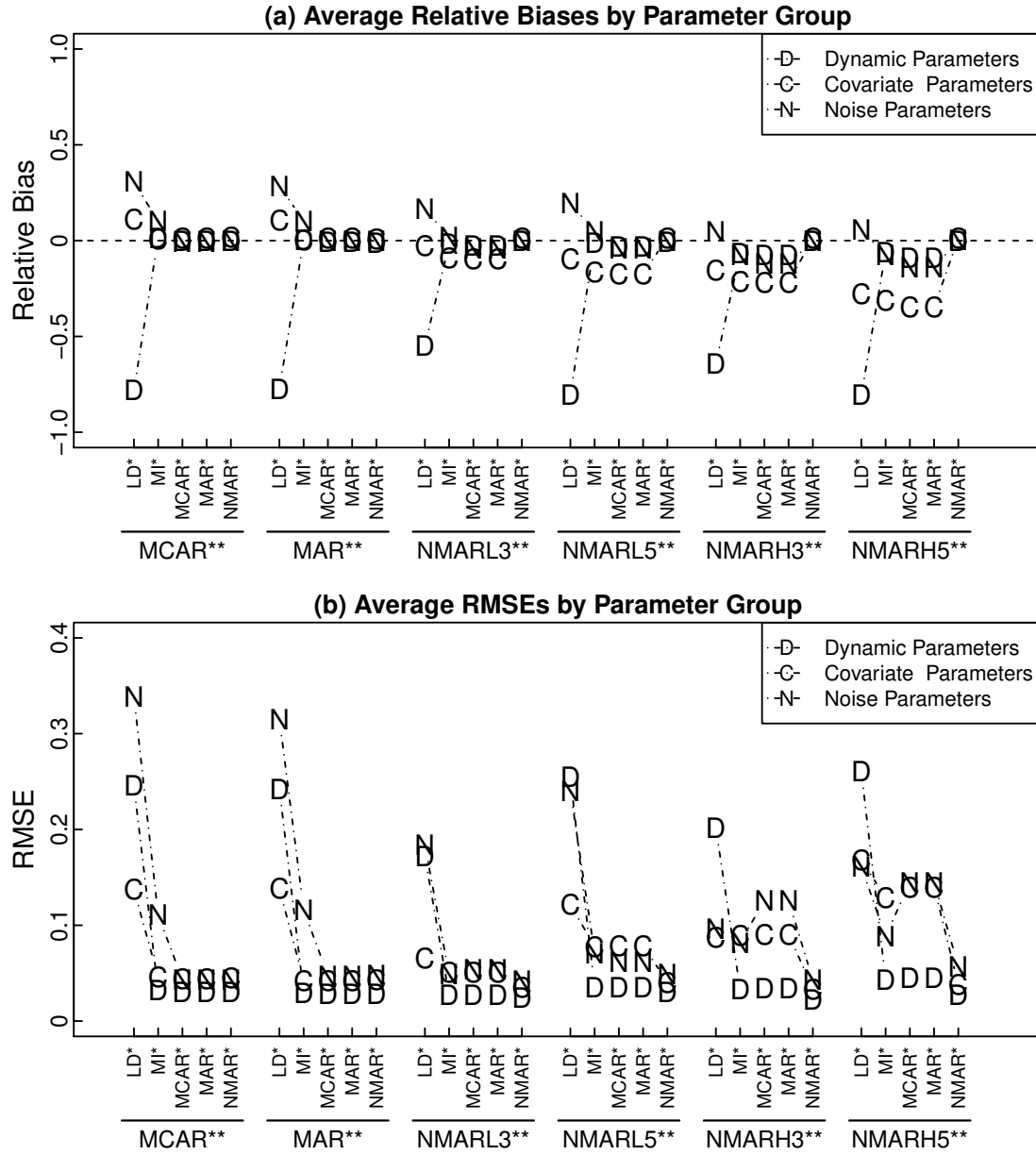


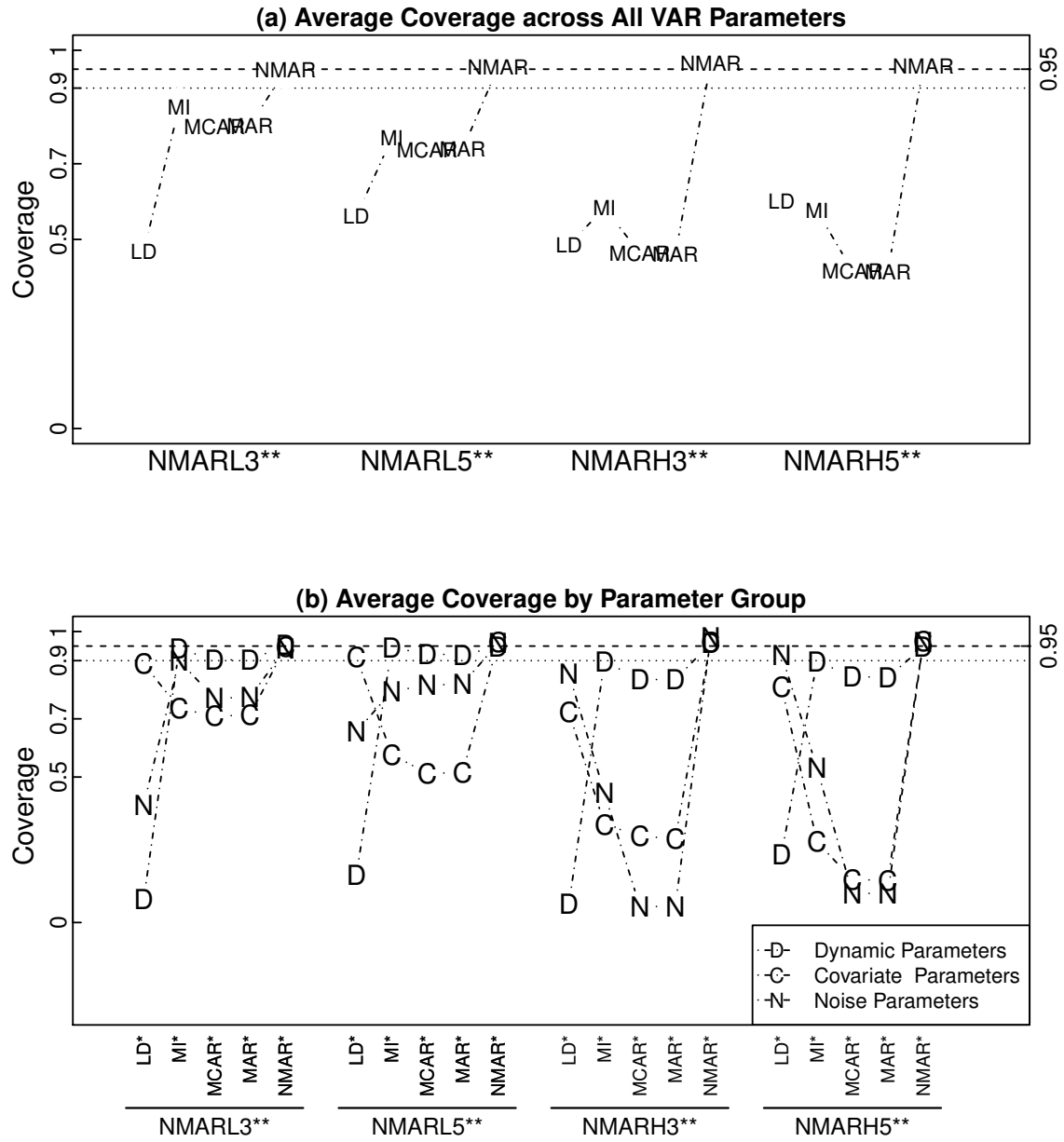
Figure 1. A Comparison between the Empirical and Simulated Distributions of Numbers of Total Observations (Ts).



** Data generation (true) model

* Method used/model fitted

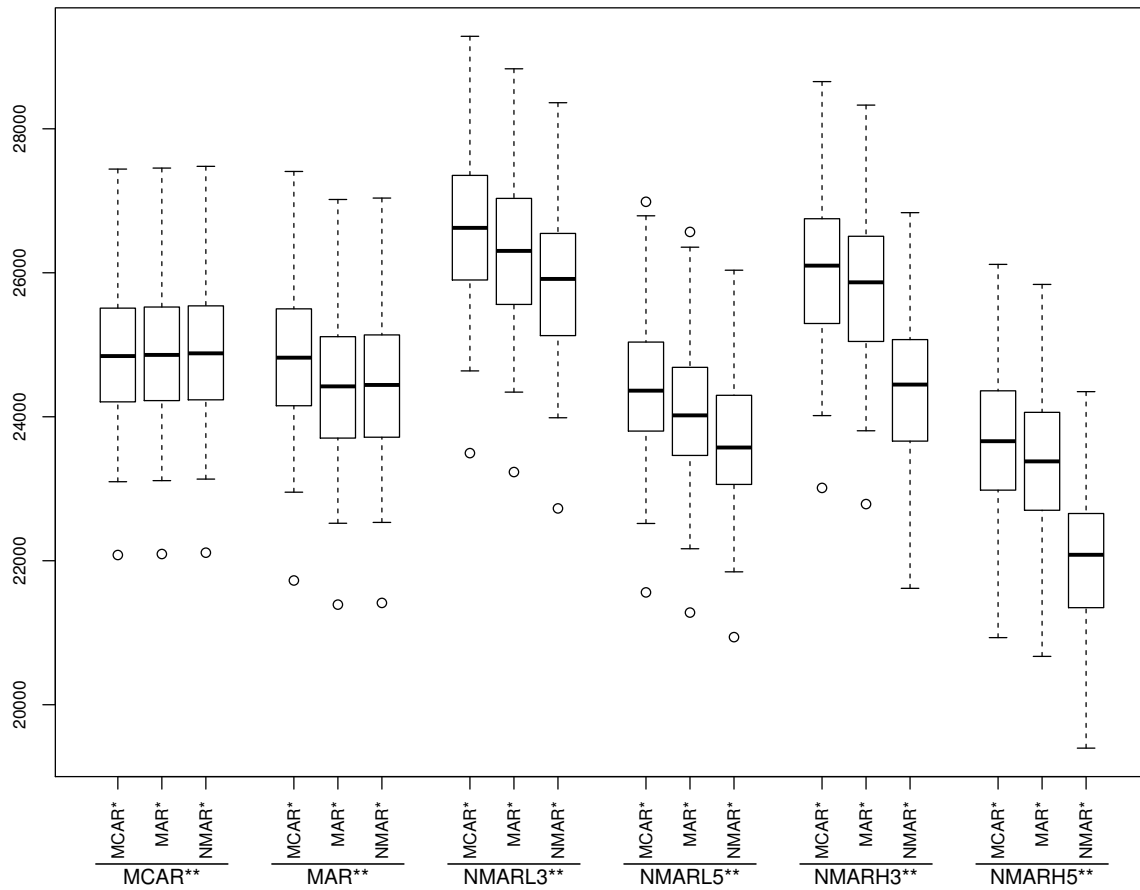
Figure 2. Relative bias (a) and RMSE (b) in parameter estimates in different approaches and when fitting true and misspecified models, grouped by parameter category.



** Data generation (true) model

* Method used/model fitted

Figure 3. Average coverage across all VAR parameters (a) and grouped by parameter categories (b) in NMAR true missingness conditions with different approaches and when fitting true and misspecified models.



** Data generation (true) model

* Fitted model

Figure 4. Boxplots of BIC values for all MC replications across true data generation conditions and missingness models fit.

Appendix A

Models for Different Missingness Mechanism for Empirical Illustration

Apart from the model specified in Equations 5 - 8, we also considered the following three alternative models for missing mechanisms, in which all the dependent variables and covariates were assumed to be MCAR, MAR, or NMAR respectively. However, they all had BIC values greater than the mixed MAR and NMAR model reported in the Empirical Illustration, indicating less good model fits (Table A1). Therefore, the mixed model (Equations 5 - 8) was selected. Here, we describe briefly these alternative missing data models considered.

Let z_{jnt} denote the j th variable in the data for family n and time t , where z_{jnt} may be any variable in \mathbf{Y} or \mathbf{X} ; and $r_{z_{jnt}}$ represents the missingness indicator for this variable.

MCAR condition: Whether an observation is missing or not does not depend on any data-related factor. Hence the missingness is modeled only by an intercept.

$$P(r_{z_{jnt}} = 1) = \text{logit}^{-1}(\phi_{int_j})$$

MAR condition: Whether an observation is missing or not depends on fully observed variables in data (here p_h and p_w).

$$P(r_{z_{jnt}} = 1 | p_{wnt}, p_{hnt}) = \text{logit}^{-1}(\phi_{int_j} + \phi_{p_w \rightarrow j} p_{wnt} + \phi_{p_h \rightarrow j} p_{hnt})$$

NMAR condition: Whether an observation is missing or not depends on fully observed variables in data as well as variables themselves.

$$P(r_{z_{jnt}} = 1 | z_{jnt}, p_{wnt}, p_{hnt}) = \text{logit}^{-1}(\phi_{int_j} + \phi_j z_{jnt} + \phi_{p_w \rightarrow j} p_{wnt} + \phi_{p_h \rightarrow j} p_{hnt})$$

Table A1

Comparison of BICs with Different Missing Mechanisms for Empirical Illustration

	Mixed*	MCAR	MAR	NMAR
BIC	14732.63	17613.94	17348.34	15447.45

*: The mixed MAR and NMAR missing mechanisms follow the specification in Equations 5 - 8

Appendix B

Alternative Model for Covariates and Results Comparison

In place of Equation 3, we considered an alternative model that contained, for x_1 :

$$x_{1nt} = \beta + \epsilon_{x_1,nt}, \quad \epsilon_{x_1,nt} \sim \mathcal{N}(0, \sigma_{\epsilon_{x_1}}^2),$$

where x_1 did not depend on any information contained in the data set.

The BIC for the alternative model was 14950.02, which was slightly higher than that of the model for covariates reported in the Empirical Illustration (Equations 3 & 4; 14732.63). Hence we selected the originally proposed model. Estimated parameter values for the alternative model are reported in Table B1. These estimates would not lead to substantively different conclusions, though some of them differ in magnitude from those in Table 1 (e.g. $d_{x_2 \rightarrow h}$).

Table B1

Estimation Result Using An Alternative Model for Covariates

Parameter	Mean	SD
VAR model parameters		
$a_{w \rightarrow w}$	0.0022	0.0391
$a_{h \rightarrow h}$	-0.0467	0.0390
$b_{h \rightarrow w}$	0.0114	0.0391
$b_{w \rightarrow h}$	0.0561	0.0391
$c_{x_1 \rightarrow w}$	-0.5112	0.0184
$c_{x_1 \rightarrow h}$	-0.5176	0.0184
$d_{x_2 \rightarrow w}$	0.2297	0.0139
$d_{x_2 \rightarrow h}$	0.2307	0.0139
σ_1^2	0.5651	0.0216
σ_2^2	0.5492	0.0211
Covariate model parameters		
β	1.7936	0.1193
ψ_1	0.7559	0.3147
ψ_2	1.5289	0.0702
ψ_3	0.4525	0.0291
ψ_4	0.0421	0.0271
$\sigma_{\epsilon_{x_1}}^2$	3.5962	0.3383
$\sigma_{\epsilon_{x_2}}^2$	3.8657	0.3471
Missingness model parameters		
$\phi_{int_{y_1}}$	-5.9098	0.4424
$\phi_{int_{y_2}}$	-5.7435	0.4095
$\phi_{int_{x_1}}$	-13.2284	2.3228
$\phi_{int_{x_2}}$	-13.2267	2.3209
$\phi_{p \rightarrow y}$	0.1822	0.1207
$\phi_{p \rightarrow x}$	-0.2788	0.1284
ϕ_x	6.0966	0.8715

Note: A number in bold indicates that the credible interval for that estimate did not contain zero.

Appendix C

R Script and JAGS Model Script Used for Empirical Illustration

R Script

```
library(rjags)
library(coda)
library(stats)
load.module("dic")

adaptSteps = 50000
burnInSteps = 10000
nChains = 2

# Setting initial values for MCMC (optional)
init1<-list(a=0,a1=0,b=0,b1=0,c=-0.5,c1=-0.5,d=0,d1=0,
            tau1=2,tau2=2,
            beta=c(0,-0.4),
            psi=c(0,0.1,0,0),
            tau_x1=0.1,
            tau_x2=0.1,
            phiy1=rep(0.1,7),
            y1=y1guess,y2=y2guess,x1=x1guess,x2=x2guess,
            .RNG.seed=2*batch+1,.RNG.name="base::Mersenne-Twister"
)

init2<-list(a=rnorm(1,0,0.1),a1=rnorm(1,0,0.1),
            b=rnorm(1,0,0.1),b1=rnorm(1,0,0.1),
            c=-0.5+rnorm(1,0,0.1),c1=-0.5+rnorm(1,0,0.1),
```

```

d=rnorm(1,0,0.1),d1=rnorm(1,0,0.1),
tau1=1,tau2=1,
beta=c(1,0.4)+rnorm(2,0,0.1),
psi=c(0.5,-0.1,0.1,0.1)+rnorm(4,0,0.1),
tau_x1=0.5,
tau_x2=0.5,
phiy1=rep(-0.1,7)+rnorm(7,0,0.1),
y1=y1guess,y2=y2guess,x1=x1guess,x2=x2guess,
.RNG.seed=2*batch,.RNG.name="base::Mersenne-Twister"
)

# Organize the data into a list
dataList <- list(y1=y1, y2=y2, x1=x1, x2=x2, N = n, T = T,
                mu0=0, tau0=0.001, R1=R1, R2=R2, R3=R3, R4=R4,
                age=age, ph=ph, pw=pw)

jagsModelNMAR <- jags.model( "emp_model.txt", data=dataList, n.chains=nChains,
                             n.adapt=adaptSteps, inits=list(init1,init2) )

update( jagsModelNMAR , n.iter=burnInSteps )

parameters<- c("a","a1","b","b1","c","c1","d","d1","Sigma1","Sigma2","deviance",
               "phiy1", "tau_x1", "tau_x2", "beta", "psi")

mcmcCodaNMAR <- coda.samples(jagsModelNMAR, variable.names=parameters,
                             n.iter=50000)

```

JAGS Script: emp_model.txt

```
model {
```

```

for (n in 1:N){

  y1[n,1] ~ dnorm(0, 0.01)
  y2[n,1] ~ dnorm(0, 0.01)
  x1[n,1] ~ dnorm( 0,0.01 )
  x2[n,1] ~ dnorm(0,0.01)

  for (t in 2:T[n]){

    mux1[n,t] <- beta[1] + beta[2]*pw[n,t] + beta[2]*ph[n,t]
    x1[n,t] ~ dnorm(mux1[n,t],tau_x1)

    mux2[n,t] <- psi[1] + psi[2]*x1[n,t] + psi[3]*pw[n,t] +
      psi[3]*ph[n,t] + psi[4]*age[n]
    x2[n,t] ~ dnorm(mux2[n,t],tau_x2)

    y1[n,t] ~ dnorm(mu1[n,t], tau1)
    y2[n,t] ~ dnorm(mu2[n,t], tau2)
    mu1[n,t]<-a*y1[n,t-1]+b*y2[n,t-1]+c*x1[n,t]+d*x2[n,t]
    mu2[n,t]<-b1*y1[n,t-1]+a1*y2[n,t-1]+c1*x1[n,t]+d1*x2[n,t]

    nmarlogity1[n,t]<-phi[1]+ phi[5]*pw[n,t] + phiy1[5]*ph[n,t]
    nmarpry1[n,t]<-exp(nmarlogity1[n,t])/(1+exp(nmarlogity1[n,t]))
    R1[n,t] ~ dbern(nmarpry1[n,t])
  }
}

```

```

nmarlogity2[n,t]<-phi[2]+ phi[5]*pw[n,t] + phi[5]*ph[n,t]
nmarpry2[n,t]<-exp(nmarlogity2[n,t])/(1+exp(nmarlogity2[n,t]))
R2[n,t] ~ dbern(nmarpry2[n,t])

nmarlogitx1[n,t]<-phi[3]+phi[7]*x2[n,t] + phi[6]*pw[n,t] +
                phi[6]*ph[n,t]
nmarprx1[n,t]<-exp(nmarlogitx1[n,t])/(1+exp(nmarlogitx1[n,t]))
R3[n,t] ~ dbern(nmarprx1[n,t])

nmarlogitx2[n,t]<-phi[4]+phi[7]*x2[n,t] + phi[6]*pw[n,t] +
                phi[6]*ph[n,t]
nmarprx2[n,t]<-exp(nmarlogitx2[n,t])/(1+exp(nmarlogitx2[n,t]))
R4[n,t] ~ dbern(nmarprx2[n,t])

}

}

# PRIORS
for (i in 1:2){
  beta[i] ~ dnorm(0, 0.01)
}
for (i in 1:4){
  psi[i] ~ dnorm(0, 0.01)
}

a ~ dnorm(0, 0.1)
a1 ~ dnorm(0, 0.1)

```

```
b    ~ dnorm(0, 0.1)
b1   ~ dnorm(0, 0.1)
c    ~ dnorm(0, 0.1)
c1   ~ dnorm(0, 0.1)
d    ~ dnorm(0, 0.1)
d1   ~ dnorm(0, 0.1)

for(i in 1:7){
  phi[i] ~ dnorm(0, 0.01)
}

tau1 ~ dgamma(0.001,0.001)
Sigma1 <- inverse(tau1)
tau2 ~ dgamma(0.001,0.001)
Sigma2 <- inverse(tau2)
tau_x1 ~ dgamma(0.001,0.001)
tau_x2 ~ dgamma(0.001,0.001)
}
```


References

- Allison, P. D. (1987). Estimation of linear models with incomplete data. *Sociological methodology*, 17(1):71–103.
- Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of abnormal psychology*, 112(4):545.
- Baltes, P. B. and Nesselroade, J. R. (1979). *Longitudinal research in the study of behavior and development*. Academic Press.
- Bolger, N. and Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. Guilford, New York.
- Borsboom, D. and Cramer, A. O. (2013). Network analysis: an integrative approach to the structure of psychopathology. *Annual review of clinical psychology*, 9:91–121.
- Bulteel, K., Ceulemans, E., Thompson, R. J., Waugh, C. E., Gotlib, I. H., Tuerlinckx, F., and Kuppens, P. (2014). Decon: A tool to detect emotional concordance in multivariate time series data of emotional responding. *Biological psychology*, 98:29–42.
- Celeux, G., Forbes, F., Robert, C. P., Titterton, D. M., et al. (2006). Deviance information criteria for missing data models. *Bayesian analysis*, 1(4):651–673.
- Chow, S.-M., Haltigan, J. D., and Messinger, D. S. (2010a). Dynamic infant-parent affect coupling during the face-to-face/still-face. *Emotion*, 10 1:101–14.
- Chow, S.-M., Hamagami, F., and Nesselroade, J. R. (2007). Age differences in dynamical cognition-emotion linkages. *Psychology and Aging*, 22(4):765–780.
- Chow, S.-M., Ho, M. R., Hamaker, E. L., and Dolan, C. V. (2010b). Equivalence and differences between structural equation modeling and state-space modeling techniques. *Structural Equation Modeling: A Multidisciplinary Journal*, 17(2):303–332.
- Chow, S.-M., Nesselroade, J. R., Shifren, K., and McArdle, J. J. (2004). Dynamic structure of emotions among individuals with Parkinson’s disease. *Structural Equation Modeling*, 11:560–582.

- Daniels, M. J. and Hogan, J. W. (2008). *Missing data in longitudinal studies: Strategies for Bayesian modeling and sensitivity analysis*. CRC Press.
- Davies, P. T. and Cummings, E. M. (1994). Marital conflict and child adjustment: An emotional security hypothesis. *Psychological bulletin*, 116(3):387.
- De Jong, P. (1991). The diffuse kalman filter. *The Annals of Statistics*, 19(2):1073–1083.
- De Silva, A. P., Moreno-Betancur, M., De Livera, A. M., Lee, K. J., and Simpson, J. A. (2017). A comparison of multiple imputation methods for handling missing values in longitudinal data in the presence of a time-varying covariate with a non-linear association with time: a simulation study. *BMC medical research methodology*, 17(1):114.
- Diggle, P. and Kenward, M. G. (1994). Informative drop-out in longitudinal data analysis. *Applied statistics*, pages 49–93.
- Ding, M., Chen, Y., and Bressler, S. L. (2006). Granger causality: basic theory and application to neuroscience. *Handbook of time series analysis: recent theoretical developments and applications*, pages 437–460.
- du Toit, S. H. and Browne, M. W. (2007). Structural equation modeling of multivariate time series. *Multivariate Behavioral Research*, 42(1):67–101.
- Emery, R. E. (1989). Family violence. *American Psychologist*, 44(2):321.
- Fahrenberg, J. and Myrtek, M. (2001). *Progress in ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies*. Hogrefe & Huber Publishers.
- Fomby, T. B., Kilian, L., and Murphy, A. (2013). *VAR Models in Macroeconomics - New Developments and Applications: Essays in Honor of Christopher A. Sims*, chapter 12, pages i–i. Emerald Insight.
- Gajewski, B. J., Lee, R., Thompson, S., Dunton, N., Becker, A., and Wells, V. (2006). Non-normal path analysis in the presence of measurement error and missing data: a bayesian analysis of nursing homes’ structure and outcomes. *Statistics in medicine*,

- 25(21):3632–3647.
- Glasser, M. (1964). Linear regression analysis with missing observations among the independent variables. *Journal of the American Statistical Association*, 59(307):834–844.
- Goeke-Morey, M. C., Papp, L. M., and Cummings, E. M. (2013). Changes in marital conflict and youths’ responses across childhood and adolescence: A test of sensitization. *Development and psychopathology*, 25(1):241–251.
- Graham, J. W. (2012). *Missing Data Analysis and Design*. Springer, New York.
- Grych, J. H. and Fincham, F. D. (1990). Marital conflict and children’s adjustment: A cognitive-contextual framework. *Psychological bulletin*, 108(2):267.
- Hamaker, E. L. (2005). Conditions for the equivalence of the autoregressive latent trajectory model and a latent growth curve model with autoregressive disturbances. *Sociological Methods & Research*, 33(3):404–416.
- Hamaker, E. L., Dolan, C. V., and Molenaar, P. C. M. (2005). Statistical modeling of the individual: Rationale and application of multivariate stationary time series analysis. *Multivariate Behavioral Research*, 40(2):207–233. PMID: 26760107.
- Harvey, A. C. (2001). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Cambridge, UK.
- Harvey, A. C. and Souza, R. C. (1987). Assessing and modelling the cyclical behaviour of rainfall in northeast Brazil. *Journal of Climate and Applied Meteorology*, 26:1317–1322.
- Heckman, J. J. (1979). Sample selection bias as a specification error (with an application to the estimation of labor supply functions). *Econometrica*, (47):153–161.
- Hemming, K. and Hutton, J. L. (2012). Bayesian sensitivity models for missing covariates in the analysis of survival data. *Journal of evaluation in clinical practice*, 18(2):238–246.
- Henderson, R., Diggle, P., and Dobson, A. (2000). Joint modelling of longitudinal

- measurements and event time data. *Biostatistics*, 1(4):465–480.
- Hogan, J. W. and Laird, N. M. (1997). Model-based approaches to analysing incomplete longitudinal and failure time data. *Statistics in medicine*, 16(3):259–272.
- Ji, L., Chow, S.-M., Schermerhorn, A. C., Jacobson, N. C., and Cummings, E. M. (2018). Handling missing data in the modeling of intensive longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, pages 1–22.
- Jones, M. P. (1996). Indicator and stratification methods for missing explanatory variables in multiple linear regression. *Journal of the American Statistical Association*, 91(433):222–230.
- Kapelner, A. and Bleich, J. (2015). Prediction with missing data via bayesian additive regression trees. *Canadian Journal of Statistics*, 43(2):224–239.
- Lee, S.-Y. (2007). *Structural Equation Modeling: A Bayesian Approach*. John Wiley & Sons.
- Little, R. J. (1993). Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association*, 88(421):125–134.
- Little, R. J. (1994). A class of pattern-mixture models for normal incomplete data. *Biometrika*, 81(3):471–483.
- Liu, S. and Molenaar, P. C. (2014). ivar: A program for imputing missing data in multivariate time series using vector autoregressive models. *Behavior research methods*, 46(4):1138–1148.
- Lu, Z. and Song, X. (2012). Finite mixture varying coefficient models for analyzing longitudinal heterogeneous data. *Statistics in medicine*, 31(6):544–560.
- Lu, Z.-H., Chow, S.-M., and Loken, E. (2017). A comparison of Bayesian and frequentist model selection methods for factor analysis models. *Psychological Methods*, 22:361–381.
- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

- Mason, A., Best, N., Richardson, S., and PLEWIS, I. (2010). Strategy for modelling non-random missing data mechanisms in observational studies using bayesian methods.
- McArdle, J. J. and Epstein, D. B. (1987). Latent growth curves within developmental structural equation models. *Child Development*, 58(1):110–133.
- Ou, L., Chow, S.-M., Ji, L., and Molenaar, P. C. (2016). (re) evaluating the implications of the autoregressive latent trajectory model through likelihood ratio tests of its initial conditions. *Multivariate Behavioral Research*, pages 1–22.
- Ou, L., Hunter, M. D., and Chow, S.-M. (2017). *dynr: Dynamic Modeling in R*. R package version 0.1.11-2.
- Plummer, M. (2016). *rjags: Bayesian Graphical Models using MCMC*. R package version 4-6.
- Plummer, M. et al. (2003). Jags: A program for analysis of bayesian graphical models using gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing*, volume 124, page 125. Vienna.
- Qin, D. (2011). Rise of VAR modelling approach. *Journal of Economic Surveys*, 25(1):156–174.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ram, N., Shiyko, M., Lunkenheimer, E. S., Doerksen, S., and Conroy, D. (2014). *Families as Coordinated Symbiotic Systems: Making use of Nonlinear Dynamic Models*, pages 19–37. Springer International Publishing, Cham.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3):581–592.
- Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American statistical Association*, 91(434):473–489.
- Schafer, J. L. (2001). *New methods for the analyses of change*, chapter Multiple imputation with PAN, pages 357–377. American Psychological Association.

- Schafer, J. L. and Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological methods*, 7(2):147.
- Scharfstein, D. O., Daniels, M. J., and Robins, J. M. (2003). Incorporating prior beliefs about selection bias into the analysis of randomized trials with missing outcomes. *Biostatistics*, 4(4):495–512.
- Schermerhorn, A. C., Chow, S.-M., and Cummings, E. M. (2010). Developmental family processes and interparental conflict: patterns of microlevel influences. *Developmental psychology*, 46(4):869–885.
- Schmittmann, V. D., Cramer, A. O., Waldorp, L. J., Epskamp, S., Kievit, R. A., and Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New ideas in psychology*, 31(1):43–53.
- Schweppe, F. (1965). Evaluation of likelihood functions for Gaussian signals. *IEEE Transactions on Information Theory*, 11:61–70.
- Schweppe, F. C. (1973). *Uncertain dynamic systems*. Upper Saddle River, NJ: Prentice Hall.
- Sinharay, S., Stern, H. S., and Russell, D. (2001). The use of multiple imputation for the analysis of missing data. *Psychological methods*, 6(4):317.
- Stone, A., Shiffman, S., Atienza, A., and Nebeling, L. (2008). *The science of real-time data capture: Self-reports in health research*. Oxford University Press, NY.
- Thomas, E. A. and Martin, J. A. (1976). Analyses of parent-infant interaction. *Psychological Review*, 83(2):141–156.
- van Buuren, S. (2012). *Flexible imputation of missing data*. CRC press.
- van Buuren, S. and Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in r. *Journal of Statistical Software*, 45(3):1–67.
- Wang, C., Daniels, M., Scharfstein, D. O., and Land, S. (2010). A bayesian shrinkage model for incomplete longitudinal binary data with application to the breast cancer prevention trial. *Journal of the American Statistical Association*, 105(492):1333–1346.

- Wang, L. and McArdle, J. J. (2008). A simulation study comparison of bayesian estimation with conventional methods for estimating unknown change points. *Structural Equation Modeling*, 15(1):52–74.
- Wu, M. C. and Carroll, R. J. (1988). Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics*, pages 175–188.
- Zhang, J., Chu, H., Hong, H., Virnig, B. A., and Carlin, B. P. (2015). Bayesian hierarchical models for network meta-analysis incorporating nonignorable missingness. *Statistical methods in medical research*, 26(5):2227–2243.
- Zheng, Y., Wiebe, R. P., Cleveland, H. H., Molenaar, P. C. M., and Harris, K. S. (2013). An idiographic examination of day-to-day patterns of substance use craving, negative affect and tobacco use among young adults in recovery. *Multivariate Behavioral Research*, 48(2):241–266.