Submitted to the Annals of Applied Statistics
arXiv: arXiv:0000.0000

BAYESIAN RANDOMIZED RESPONSE TECHNIQUE WITH MULTIPLE SENSITIVE ATTRIBUTES: THE CASE OF INFORMATION SYSTEMS RESOURCE MISUSE

By Ray S. W. Chung[†], Amanda M. Y. Chu[‡] and Mike K. P. So[*,†]

The Hong Kong University of Science and Technology[†] and Hang Seng Management College[‡]

The randomized response technique (RRT) is a classical and effective method used to mitigate the distortion arising from dishonest answers. The traditional RRT usually focuses on the case of a single sensitive attribute, and discussion of the case of multiple sensitive attributes is limited. Here, we study a business case to identify some individual and organizational determinants driving information systems (IS) resource misuse in the workplace. People who actually engage in IS resource misuse are probably not willing to provide honest answers, given the sensitivity of the topic. Yet, to develop the causal relationship between IS resource misuse and its determinants, a version of the RRT for multivariate analysis is required. To implement the RRT with multiple sensitive attributes, we propose a Bayesian approach for estimating covariance matrices with incomplete information (resulting from the randomization procedure in the RRT case). The proposed approach (i) accommodates the positive definite condition and other intrinsic parameter constraints in the posterior to improve statistical precision, (ii) incorporates Bayesian shrinkage estimation for covariance matrices despite incomplete information, and (iii) adopts a quasi-likelihood method to achieve Bayesian semiparametric inference for enhancing flexibility. We show the effectiveness of the proposed method in a simulation study. We also apply the Bayesian RRT method and structural equation modeling to identify the causal relationship between IS resource misuse and its determinants.

1. Introduction.

Surveys involving multiple sensitive questions appear in various fields of applications (Minsky-Kelly et al., 2005; Nuno and St John, 2015; Höglinger et al., 2016; Rosenfeld et al., 2016). Existing statistical in-

[*] The authors thank the editor, associate editor, and two reviewers for helpful suggestions and constructive comments. This paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (UGC/FDS14/P05/16) and the Hong Kong RGC General Research Fund (Grant number 604313).

MSC 2010 subject classifications: Primary 62D05

Keywords and phrases: Causal modeling, Markov chain Monte Carlo, quasi-likelihood, sensitive responses, shrinkage estimation of covariance matrix, unrelated question design

ference techniques mainly focus on analyzing a single sensitive question or bivariate relationship. In this paper, we address a research gap by proposing a Bayesian method to estimate the covariance of multiple sensitive and direct questions in surveys while ensuring positive definiteness and allowing possible sparsity in the covariance matrix of all responses of interest. We use a business case on information systems (IS) resource misuse to illustrate our Bayesian method. IS resource misuse is defined as the unauthorized use of any IS resources, including applications, the Internet, and networks in the workplace (Chu et al., 2015, p. 521). It is a serious problem in the workplace because it may increase the risk of attacks on organizations' computer systems in addition to company information loss or leakage (Lin and Ding, 2003; D'Arcy and Devaraj, 2012). We launch a survey with questions measuring the factors of interest and actual IS resource misuse behavior. However, as IS resource misuse is unpleasant behavior that may incur legal liability, respondents who actually misuse IS resources would probably not provide honest answers to those questions (Locander et al., 1976).

Warner (1965) introduced a randomization scheme, usually known as the randomized response technique (RRT), to mitigate the response distortion resulting from sensitive questions. Under RRT settings, respondents are asked either "whether you have a sensitive attribute" or "whether you don't have a sensitive attribute". Assigning of these questions is randomized to ensure respondents' privacy, but the researchers can still estimate the proportion of the population with the sensitive attribute from the randomized responses. Numerous extensions and alternative designs of Warner's RRT have been developed in the past 50 years (e.g., Kuk, 1990; Mangat and Singh, 1990; Mangat, 1994; Gjestvang and Singh, 2006). In particular, Horvitz et al. (1967) and Greenberg et al. (1969) proposed the unrelated question design (UQD), where respondents are instructed to answer a sensitive question or an unrelated and innocuous question at random. The UQD has an advantage over other designs in that it can also handle quantitative data (Greenberg et al., 1971; Pollock and Bek, 1976). There are also other designs of the RRT, including the forced response design, the disguised response design (Blair et al., 2015), the randomized sum score/list experiment (Raghavarao and Federer, 1979; Cruyff et al., 2008; Imai, 2011; Blair and Imai, 2012; Imai et al., 2015; Rosenfeld et al., 2016), and the nonrandomized response techniques (Tan et al., 2009) adopted in various domains. Comprehensive reviews of the RRT can be found in Fox and Tracy (1986), Singh (2003) and Chaudhuri (2011).

In the first 20 years of RRT research after Warner (1965), works focused on collecting and analyzing a single sensitive attribute, i.e., asking one sensitive question. However, our study estimates the dependency in a set of sensitive attributes, which are asked under a randomization procedure, and a set of standard attributes, which are asked directly. Discussions of the RRT in this kind of setting are very limited in the literature. Some exceptions are Tamhane (1981) and Christofides (2005), who extended the method to deal with more than one sensitive dichotomous question. Yet, their results cannot be applied to our study, as the questions in our survey are quantitative. Fox and Tracy (1984) investigated a bivariate case and estimated the correlation between two sensitive responses under the UQD, with the unrelated and innocuous question distribution assumed to be known. Kwan et al. (2010) handled a more general case under the UQD to capture the covariance between a randomized sensitive response and a direct response. Nevertheless, both Fox and Tracy (1984) and Kwan et al. (2010) ignored the intrinsic positive definite constraint in a covariance matrix, and hence we can easily produce invalid covariance matrices (see Section 4) by their methods. As several questions are included in a wide range of surveys, the problem of invalid covariance matrices is highly significant if we apply the method of Fox and Tracy (1984) or Kwan et al. (2010).

We restrict our discussion to the UQD in our study, as the questions in our survey are quantitative rather than the typical binary "Yes"/"No" response of most of other RRT designs. To deal with the problem of invalid estimates of covariance matrices, we develop a Bayesian RRT method that is applied specifically to the UQD. Applications of the Bayesian approach to RRT problems have been discussed in the literature. For example, Chen and Singh (2011) and Jayaraj et al. (2014) studied pseudo-Bayes estimators for one sensitive dichotomous question. Blair and Zhou (2016) adopted a Bayesian implementation of the method in Blair et al. (2015). In contrast, our Bayesian RRT method is designed to return a positive-definite estimate of the covariance matrix of a set of randomized sensitive and direct questions simultaneously, even under a high-dimensional setting. In our Bayesian RRT approach, we reparameterize the covariance matrix by the modified Cholesky decomposition to ensure its positive definiteness. We use a Laplace prior to impose sparsity on the inference of the covariance matrix for enhancing statistical precision. In addition, we adopt a quasi-likelihood constructed from moment equations to allow the distribution of attributes to be unspecified, thus enhancing the flexibility of the RRT modeling. We also prove the posterior consistency to justify the use of the quasi-likelihood and show that the Bayesian RRT method can increase the stability of the covariance-matrix estimates even when the dimension of variables is high.

In Section 2, we describe the UQD and the challenges in enforcing the positive definiteness constraint when estimating the covariance matrix with multiple sensitive attributes. Section 3 presents the Bayesian RRT method with the quasi-likelihood. In Section 4, we conduct a simulation study to compare the performance of the Bayesian RRT method with a benchmark method-of-moments estimator, and highlight the superiority of the Bayesian approach. In Section 5, we apply the Bayesian RRT method to investigate the individual and organizational determinants triggering IS resource misuse in the workplace, and provide a detailed explanation of the research model and the steps of the survey. Section 6 presents the conclusion and discussion.

2. Unrelated question design.

2.1. *Single sensitive attribute setting.* Horvitz et al. (1967) and Greenberg et al. (1971) proposed the UQD to decrease the distortion of responses to quantitative questions in a survey when the questions are sensitive to respondents. Under the UQD, respondents are presented with a pair of questions. One question is sensitive and of interest, while the other question is innocuous, unrelated to the sensitive question, and not of interest. In statistical terms, the unrelated question response is assumed to be independent of the sensitive question response. The respondents are instructed to generate a dichotomous outcome from a private random choice procedure, and answer only one of the two questions according to the generated outcome. For example, the respondent may be asked to draw a card from a poker deck and keep the card secret. The respondent then answers the sensitive question if the card drawn is red or the unrelated question if it is not red, without disclosing which question he or she actually answers. As the outcome from the private random choice procedure is unknown to the interviewer, the interviewer does not know which question has been answered. In this way, the true response of the respondent to the sensitive question is masked by the private random choice procedure. It is hoped that this encourages the respondent to give a truthful response. In our experimental setting, we require that the probability governing the private random choice procedure be known. In addition, we need to collect two samples, named Samples 1 and 2, and assign different probabilities governing the private random choice procedure to the two samples. Greenberg et al. (1969) demonstrated that it is not necessary to split the sample into two pieces provided that the mean of the unrelated question response is known in advance.

2.2. *Multiple sensitive attribute setting.* Assume there are p sensitive questions, and that each sensitive question is paired with an unrelated question. Respondents are asked to generate a random outcome for each pair of sensitive and unrelated questions to determine which question is answered for that pair. We require that the random outcomes of the question pairs be independent, and that the probability of answering sensitive questions be kept constant for all of the question pairs within a particular sample. Let $S_1, \ldots, S_p$ be the responses to the p sensitive questions and $U_1, \cdots, U_p$ be the responses to the p unrelated questions. Let n_k be the size of Sample k, ω_k be the probability of respondents in Sample k answering the sensitive question in any question pair, and $I_{ki} \overset{\text{iid}}{\sim} \text{Bernoulli}(\omega_k)$ be the binary variables indicating whether a respondent in Sample k answers the ith sensitive ($I_{ki} = 1$) or the ith unrelated question ($I_{ki} = 0$), for $k = 1, 2$. For Z_{ki} denoting the observed response of a respondent in Sample k to the ith randomized question, it can be expressed as $Z_{ki} = I_{ki}S_i + (1 - I_{ki})U_i$, for $i = 1, \ldots, p$ and $k = 1, 2$. Under the UQD, we assume (i) ω_k is known, (ii) Samples 1 and 2 are homogeneous, and (iii) $\omega_1 \neq \omega_2$.

Here, $S_1, \ldots, S_p$ are measured using the RRT because they are sensitive. Assume that there are also q attributes of interest that are not sensitive and their responses $D_1, \ldots, D_q$ can be measured by direct questioning. Let $\mathbf{S} = (S_1, \ldots, S_p)^T$ be the vector of sensitive question responses, $\mathbf{U} = (U_1, \ldots, U_p)^T$ be the vector of unrelated question responses, $\mathbf{D} = (D_1, \ldots, D_q)^T$ be the vector of direct question responses, and $\mathbf{Y} = (\mathbf{S}^T, \mathbf{D}^T)^T$ be the pooled responses of interest. Define $\mathbf{a}_p$ to be a pth dimensional vector and $\mathbf{a}_{p\times q}$ to be a $p \times q$ matrix, both with all entries equal to a, for any real number a. The $(p+q)\times 1$ vector of the pooled observed responses in Sample k for $k = 1, 2$, denoted by $\mathbf{Z}_k$, is

$$\mathbf{Z}_k = \begin{pmatrix} \mathbf{I}_k \odot \mathbf{S} + (\mathbf{1}_p - \mathbf{I}_k) \odot \mathbf{U} \\ \mathbf{D} \end{pmatrix}, \tag{2.1}$$

where $\odot$ is the pointwise (Hadamard) product and $\mathbf{I}_k = (I_{k1}, \ldots, I_{kp})^T$.

As $\mathbf{U}$ reflects the responses to the unrelated questions, $\mathbf{Y}$ and $\mathbf{U}$ are assumed to be independent. To enable wide applicability and to accommodate various types of survey responses, we leave the distribution of $\mathbf{Y}$ and $\mathbf{U}$ unspecified. In conducting multivariate analysis, we are interested in the covariance matrix of the responses to the questions of interest $\mathbf{Y}$, denoted by $\mathbf{\Sigma}_Y$. Furthermore, let $\mathbf{\Sigma}_U$ be the covariance matrix of $\mathbf{U}$. Without loss of generality, we assume $E(\mathbf{Y}) = \mathbf{0}_{p+q}$ and $E(\mathbf{U}) = \mathbf{0}_p$. By the derivation given in Appendix A,

$$\mathbf{\Sigma}_Y = \mathbf{C}_{1*} \odot E(\mathbf{Z}_1\mathbf{Z}_1^T) + \mathbf{C}_{2*} \odot E(\mathbf{Z}_2\mathbf{Z}_2^T), \tag{2.2}$$

where $\mathbf{C}_{k*} = \mathbf{C}_0^{\circ -1} \odot \mathbf{C}_k$,

$$\mathbf{C}_1 = \begin{pmatrix} \boldsymbol{\mathcal{P}}_2^c & \mathbf{1}_{p\times q} \\ \mathbf{1}_{q\times p} & \mathbf{1}_{q\times q} \end{pmatrix}, \quad \mathbf{C}_2 = \begin{pmatrix} -\boldsymbol{\mathcal{P}}_1^c & \mathbf{1}_{p\times q} \\ \mathbf{1}_{q\times p} & \mathbf{1}_{q\times q} \end{pmatrix},$$

$$\mathbf{C}_0 = \begin{pmatrix} \boldsymbol{\mathcal{P}}_2^c \odot \boldsymbol{\mathcal{P}}_1 - \boldsymbol{\mathcal{P}}_1^c \odot \boldsymbol{\mathcal{P}}_2 & (\omega_{\mathbf{1}} + \omega_{\mathbf{2}})_{p\times q} \\ (\omega_{\mathbf{1}} + \omega_{\mathbf{2}})_{q\times p} & \mathbf{2}_{q\times q} \end{pmatrix},$$

$\mathbf{A}^{\circ -1}$ denotes the Hadamard inverse of a matrix $\mathbf{A}$[1], $\boldsymbol{\mathcal{P}}_k$ is a $p \times p$ matrix with diagonal elements given by ω_k and other elements given by p_k^2, $\boldsymbol{\mathcal{P}}_k^c$ is defined in the same way as $\boldsymbol{\mathcal{P}}_k$ but with ω_k replaced by $1 - \omega_k$, for $k = 1, 2$, and $(\omega_{\mathbf{1}} + \omega_{\mathbf{2}})_{p\times q}$ is a $p \times q$ matrix with all entries equal to $\omega_1 + \omega_2$. As in the cases where we have a single sensitive attribute, only one sample is needed if $\mathbf{\Sigma}_U$ is known. However, in the rest of the discussion, we assume that $\mathbf{\Sigma}_U$ of the unrelated questions is unknown, making two samples necessary.

In theory, we can apply the method of moments and (2.2) to estimate $\mathbf{\Sigma}_Y$. However, the resulting estimator suffers from two problems: (i) the estimate of $\mathbf{\Sigma}_Y$ may not be positive definite and (ii) the estimation error will increase drastically when the dimension of $\mathbf{Y}$ grows. To handle these problems, we introduce a Bayesian RRT method that is designed for the multiple-attribute UQD problem.

3. A Bayesian RRT method.

3.1. *Reparameterization.* Clumsy constraints, such as the smallest eigenvalue being greater than 0, are usually imposed on the space of the prior distribution of $\mathbf{\Sigma}_Y$ to ensure its positive definiteness. This makes the inference complicated, especially when the dimension of $\mathbf{\Sigma}_Y$ is high. To avoid nontrivial and clumsy constraints, we reparameterize $\mathbf{\Sigma}_Y$ based on a modified Cholesky decomposition. Following the idea of Huang et al. (2006), we interpret the modified Cholesky decomposition as a series of linear regressions. $\mathbf{Y}$ is expressed as

$$Y_i = \begin{cases} \varepsilon_i & \text{if } i = 1, \\ \sum_{j=1}^{i-1} \phi_{i,j} Y_j + \varepsilon_i & \text{if } 2 \le i \le p+q, \end{cases} \tag{3.1}$$

where the ε_i's are iid errors with variance v_i. Therefore,

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{p+q} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ \phi_{2,1} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \phi_{p+q,1} & \phi_{p+q,2} & \phi_{p+q,3} & \cdots & 0 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_{p+q} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_{p+q} \end{pmatrix}, \tag{3.2}$$

[1]If $a_{i,j}$ is the (i,j)th element of $\mathbf{A}$, the (i,j)th element of $\mathbf{A}^{\circ -1}$ is given by $a_{i,j}^{-1}$.

which implies $\boldsymbol{\varepsilon} = \mathbf{TY}$, where $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_{p+q})^T$ and $\mathbf{T}$ is a lower triangular matrix with diagonal elements equal to 1 and the (i, j)th element equal to $-\phi_{i,j}$ for $i > j$. Using the covariance of $\boldsymbol{\varepsilon} = \mathbf{TY}$, we can recover $\mathbf{\Sigma}_Y$ from the equation

$$\mathbf{T}\mathbf{\Sigma}_Y\mathbf{T}^T = \mathbf{B}, \tag{3.3}$$

where $\mathbf{B}$ is a diagonal matrix with its diagonal given by $(v_1, \ldots, v_{p+q})^T$. After reparameterization by (3.3), the parameters of interest switch from $\{\sigma_{i,j}\}$, the elements of $\mathbf{\Sigma}_Y$, to $\{\phi_{i,j}\}$ and $\{v_i\}$. We also reparameterize $\{v_i\}$ to $\{\log v_i\}$. Under the new parameterization, the spaces of $\phi_{i,j}$ and $\log v_i$ are on the whole real line, without further constraints imposed on the parameter space. Thus, the posterior distribution of $\mathbf{\Sigma}_Y$ is naturally supported in the space of positive definite matrices. In this way, we avoid complicated support in the prior distribution of $\mathbf{\Sigma}_Y$.

3.2. *Bayesian inference with quasi-likelihood.* After the reparameterization by (3.3), the inference on $\mathbf{\Sigma}_Y$ is made through the $\phi_{i,j}$'s and $\log v_i$'s. We introduce the Bayesian inference on the $\phi_{i,j}$'s and $\log v_i$'s in this section.

3.2.1. *Prior distribution to induce sparsity.* We assume that the prior distributions of the $\phi_{i,j}$'s and $\log v_i$'s are independent. When $\mathbf{Y}$ is of high dimension, it may be that many of the variables in $\mathbf{Y}$ are pairwise-independent or the dependence is negligible. In this case, the off-diagonal elements of $\mathbf{\Sigma}_Y$ may contain many zero values. This feature is usually known as covariance sparsity. Referring to (3.1), the sparsity of $\mathbf{\Sigma}_Y$ can be interpreted as the existence of several zero values in $\{\phi_{i,j}\}$. In addition, (3.1) gives us insight that $\phi_{i,j}$ behaves like a regression parameter. Motivated by Park and Casella (2008), who imposed shrinkage on regression parameters through the Laplace prior, we specify the prior distribution of $\phi_{i,j}$ as a Laplace distribution $\pi(\phi_{i,j}) \propto \exp(-\lambda|\phi_{i,j}|)$, where λ is a nonnegative hyperparameter and a larger λ represents a stronger shrinkage effect on $\phi_{i,j}$. Given that sparsity exists, Park and Casella (2008) demonstrated that the Laplace prior could effectively decrease the estimation error. We illustrate the advantage of the shrinkage prior in Section 4. We set the prior distribution of $\log v_i$ to be normal with mean 0 and variance η.

3.2.2. *Quasi-likelihood.* We proceed to define a quasi-likelihood of $\mathbf{Y}$ for the Bayesian inference. In practice, the entries of $\mathbf{Y}$ may possess very different statistical properties. For example, the first variable of $\mathbf{Y}$ may be discrete and the second variable of $\mathbf{Y}$ may be continuous. As such, specifying the joint distribution of $\mathbf{Y}$ is often difficult. However, the traditional

Bayesian inference cannot be conducted with likelihoods omitted. In this paper, we follow the approach of Yin (2009) to construct a quasi-likelihood from generalized method of moments (GMM) equations. Let $\mathbf{Z}_{kt}$ be an iid copy of $\mathbf{Z}_k$ representing the observed responses of the tth respondent in Sample k for $k = 1, 2$. Assume that $\boldsymbol{\theta}$ is the h-dimensional parameter vector collecting $\phi_{i,j}$'s and $\log v_i$'s, h is the number of parameters given by $(p + q)(p + q + 1)/2$, and $\boldsymbol{\Sigma}_Y(\boldsymbol{\theta})$ is $\boldsymbol{\Sigma}_Y$ but reparameterized in terms of $\boldsymbol{\theta}$ through (3.3). In addition, assume that $n = \min(n_1, n_2) \to \infty$, for $k = 1, 2$. According to (2.2), we can form a moment equation given by $\mathbf{G}_n(\boldsymbol{\theta}) = \mathbf{0}$, where

(3.4)

$$\mathbf{G}_n(\boldsymbol{\theta}) = \text{vech}\left(n_1^{-1}\mathbf{C}_{1*} \odot \sum_{t=1}^{n_1} \mathbf{Z}_{1t}\mathbf{Z}_{1t}^T + n_2^{-1}\mathbf{C}_{2*} \odot \sum_{t=1}^{n_2} \mathbf{Z}_{2t}\mathbf{Z}_{2t}^T - \boldsymbol{\Sigma}_Y(\boldsymbol{\theta})\right).$$

Some variables of $\mathbf{G}_n(\boldsymbol{\theta})$ may be more important than others; hence, we need to allocate higher weights to those variables. Therefore, we define the weighted moment equation as $\mathbf{G}_n^{\mathcal{W}}(\boldsymbol{\theta}) = \mathbf{0}$, where $\mathbf{G}_n^{\mathcal{W}}(\boldsymbol{\theta}) = \mathbf{W}\mathbf{G}_n(\boldsymbol{\theta})$ and $\mathbf{W}$ is a diagonal matrix with its diagonal elements equal to the weights of the corresponding variables of $\mathbf{G}_n(\boldsymbol{\theta})$. Note that each variable of $\mathbf{G}_n(\boldsymbol{\theta})$ is the moment equation of a particular entry of $\boldsymbol{\Sigma}_Y$. In addition, the complexity of the variable of $\mathbf{G}_n(\boldsymbol{\theta})$ corresponding to $\sigma_{i,j}$, $i > j$, increases with i, as the variable with larger i contains more elements of $\boldsymbol{\theta}$. Based on our empirical experience, the accuracy will be increased if the more complex variable of $\mathbf{G}_n(\boldsymbol{\theta})$ is down-weighted. Therefore, the variable of $\mathbf{G}_n(\boldsymbol{\theta})$ corresponding to $\sigma_{i,j}$, $i > j$, is assigned a weight of $w - i$, where $w = p + q$ is the dimension of $\mathbf{Y}$ in our settings.

From (3.4) and the definition of $\mathbf{G}_n^{\mathcal{W}}(\boldsymbol{\theta})$, $\mathbf{V}_n =$ var$(\mathbf{G}_n^{\mathcal{W}}(\boldsymbol{\theta}))$ can be estimated by $\widehat{\mathbf{V}}_n = \mathbf{W}\left\{n_1^{-2}\sum_{t=1}^{n_1}(\mathbf{g}_{1t} - \overline{\mathbf{g}}_1)(\mathbf{g}_{1t} - \overline{\mathbf{g}}_1)^T + n_2^{-2}\sum_{t=1}^{n_2}(\mathbf{g}_{2t} - \overline{\mathbf{g}}_2)(\mathbf{g}_{2t} - \overline{\mathbf{g}}_2)^T\right\}\mathbf{W}$, where $\mathbf{g}_{kt} = \text{vech}(\mathbf{C}_{k*} \odot (\mathbf{Z}_{kt}\mathbf{Z}_{kt}^T))$ and $\overline{\mathbf{g}}_k = n_k^{-1}\sum_{t=1}^{n_k}\mathbf{g}_{kt}$ for $k = 1, 2$. We define the GMM estimator $\widehat{\boldsymbol{\theta}}_n$ as the vector minimizing the quadratic objective function $Q_n(\boldsymbol{\theta}) = (\mathbf{G}_n^{\mathcal{W}}(\boldsymbol{\theta}))^T\widehat{\mathbf{V}}_n^{-1}\mathbf{G}_n^{\mathcal{W}}(\boldsymbol{\theta})$. The asymptotic distribution of $\widehat{\boldsymbol{\theta}}_n$ is given by $(\mathbf{H}_n^T\mathbf{V}_n^{-1}\mathbf{H}_n)^{1/2}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, \mathbf{I})$, where $\boldsymbol{\theta}_0$ is the true value of $\boldsymbol{\theta}$, $\mathbf{H}_n = E(\dot{\mathbf{G}}_n^{\mathcal{W}}(\boldsymbol{\theta}_0))$, and $\dot{\mathbf{f}} = \frac{\partial}{\partial\boldsymbol{\theta}}\mathbf{f}(\boldsymbol{\theta})$ for $\mathbf{f}(\boldsymbol{\theta})$ being a function of $\boldsymbol{\theta}$ (Hansen, 1982). Yin (2009) considered $\exp(-Q_n(\boldsymbol{\theta})/2)$ as the quasi-likelihood for the Bayesian inference. Yin (2009) justified the use of the quasi-likelihood based on the intuition that the quasi-likelihood and the true likelihood have the same asymptotic behavior. To strengthen the argument of Yin (2009), we introduce the following theorem, and give its proof in Appendix B.

THEOREM 3.1. *Suppose that*

1. *there exists a sequence of GMM estimators* $\{\widehat{\boldsymbol{\theta}}_n\}$ *s.t.* $P(\boldsymbol{G}_n^{\mathcal{W}}(\widehat{\boldsymbol{\theta}}_n) = \mathbf{0}) \longrightarrow 1$, $\widehat{\boldsymbol{\theta}}_n \xrightarrow{p} \boldsymbol{\theta}_0$ *as* $n \to \infty$,
2. $\mathbf{V}_n$ *and* $\mathbf{H}_n$ *are invertible for large* n,
3. *for any* $\varepsilon > 0$, *there exists* $\delta > 0$ *s.t.*

$$P\left\{\sup_{\boldsymbol{\theta}\in N_0(\delta)} \|[\dot{\mathbf{G}}_n^{\mathcal{W}}(\boldsymbol{\theta}_0)]^{-1}(\dot{\mathbf{G}}_n^{\mathcal{W}}(\boldsymbol{\theta}) - \dot{\mathbf{G}}_n^{\mathcal{W}}(\boldsymbol{\theta}_0))\| > \varepsilon\right\} \longrightarrow 0,$$

   where $N_0(\delta)$ *is the neighborhood of* $\boldsymbol{\theta}_0$ *with radius of* δ,
4. $\widehat{\mathbf{V}}_n^{-1}\mathbf{V}_n \xrightarrow{p} \mathbf{I}$, $(\dot{\mathbf{G}}_n^{\mathcal{W}}(\boldsymbol{\theta}_0))^{-1}\mathbf{H}_n \xrightarrow{p} \mathbf{I}$,
5. $\|(\mathbf{H}_n^T\mathbf{V}_n^{-1}\mathbf{H}_n)^{-1}\| = o(1)$ *and*
6. $\pi(\cdot)$ *is independent of* n, *continuous at* $\boldsymbol{\theta}_0$, *and positive at* $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

For $\widehat{\pi}(\boldsymbol{\theta} \mid \mathbf{z}, \mathbf{d}) \propto \pi(\boldsymbol{\theta}) \exp(-1/2 \cdot Q_n(\boldsymbol{\theta}))$ *denoting the quasi-posterior density of* $\boldsymbol{\theta}$ *and* $\widehat{\pi}(\boldsymbol{\eta} \mid \mathbf{z}, \mathbf{d})$ *denoting the quasi-posterior density of* $\boldsymbol{\eta} = (\mathbf{H}_n^T\mathbf{V}_n^{-1}\mathbf{H}_n)^{1/2}(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_n)$, $\widehat{\pi}(\boldsymbol{\eta} \mid \mathbf{z}, \mathbf{d}) \xrightarrow{p} \phi(\boldsymbol{\eta})$ *for any* $\boldsymbol{\eta} \in (-\infty, \infty)^d$, *where* $\phi(\boldsymbol{\eta}) \propto \exp(-\|\boldsymbol{\eta}\|^2/2)$.

As the dimensions of $\boldsymbol{G}_n^{\mathcal{W}}(\boldsymbol{\theta})$ and $\boldsymbol{\theta}$ are identical in our settings, we may find $\widehat{\boldsymbol{\theta}}_n$ that solves the equation $\boldsymbol{G}_n^{\mathcal{W}}(\widehat{\boldsymbol{\theta}}_n) = \mathbf{0}$ exactly. In such a case, $\widehat{\boldsymbol{\theta}}_n$ can be regarded as a generalized estimating equation (GEE) estimator (Shao, 2003, Chapter 5.4) in addition to the GMM estimator. Shao (2003, Proposition 5.3) suggested standard sufficient conditions for condition 1 in Theorem 3.1 under the GEE framework. Conditions 2-5 in Theorem 3.1 are also variations of the standard sufficient conditions for the asymptotic normality of the GEE estimator (Shao, 2003, Theorem 5.14). Theorem 3.1 implies that with a large sample size, the quasi-posterior distribution of $\boldsymbol{\theta}$ behaves like the normal distribution with mean $\widehat{\boldsymbol{\theta}}_n$ and covariance $(\mathbf{H}_n^T\mathbf{V}_n^{-1}\mathbf{H}_n)^{-1}$. Remember that the asymptotic distribution of $\widehat{\boldsymbol{\theta}}_n$ under the frequentist approach is given by the multivariate normal with mean $\boldsymbol{\theta}_0$ and covariance $(\mathbf{H}_n^T\mathbf{V}_n^{-1}\mathbf{H}_n)^{-1}$. Therefore, provided that the regularity conditions of Theorem 3.1 are satisfied, the limiting quasi-posterior distribution provides us with information analogous to that provided by the limiting sampling distribution of $\widehat{\boldsymbol{\theta}}_n$. This justifies the validity of the information provided by the quasi-posterior distribution.

3.3. *MCMC sampling algorithm.* As the quasi-posterior distribution $\widehat{\pi}(\boldsymbol{\theta} \mid \mathbf{z}, \mathbf{d})$ is analytically intractable, we adopt Markov chain Monte Carlo (MCMC) sampling with a Metropolis–Hastings (MH) algorithm to draw samples from $\widehat{\pi}(\boldsymbol{\theta} \mid \mathbf{z}, \mathbf{d})$. Nevertheless, the traditional MH algorithm with random-walk proposal kernel fails to work, as the dimension of $\boldsymbol{\theta}$ is very high: over 100 in our case. Thus, the resulting curse of dimensionality and multi-modal nature severely hinder the sampling efficiency. To tackle the problem, we adopt a multiple-try algorithm (Liu et al., 2000). The multiple-try algorithm is designed to propose multiple trial points in each sampling iteration, increasing the chances of acceptance and of escaping from local modes. This facilitates the efficient sampling of $\boldsymbol{\theta}$ with a high dimension in our RRT cases.

We block $\boldsymbol{\theta}$ as follows during the sampling. $\{\phi_{i,j}\}$ and $\{\log v_i\}$ are blocked on a row-by-row basis in relation to the matrices shown in (3.2). In other words, each block is a subvector of $\boldsymbol{\theta}$ in the form $(\phi_{i,1}, \ldots, \phi_{i,i-1}, \log v_i)$, $i = 1, \ldots, p+q$. If the number of elements within a particular block is too large, we subdivide the block into smaller blocks by (i) permuting the elements within the block and (ii) separating the block into smaller blocks using the "stochastic knots" method (Shephard and Pitt, 1997). From our experience in both simulated and real RRT examples, the MCMC algorithm proposed facilitates efficient sampling of $\boldsymbol{\theta}$ with a fast mixing rate.

The MCMC simulation is coded in R and executed on a Linux platform installed in a computer with a 3.50 GHz Intel Core i7-4771 CPU and 32 GB memory. The R function `BayeRRT()`, which carries out the MCMC simulation in this section, can be accessed at XXXXX. We can follow the instruction of the HTML file attached to `BayeRRT()` to obtain the results of one simulated dataset as an example. `BayeRRT()` is also capable of implementing an adaptation algorithm that further increases the mixing rate of the Markov chain. Section 5.3 presents details of the adaptation algorithm.

4. Simulation study.

4.1. *Design of the simulation study.* In this section, we outline the simulation study to investigate the performance of our proposed Bayesian estimator. We consider two $\boldsymbol{\Sigma}_Y$ to give two dependence structures of $\boldsymbol{Y}$. The first $\boldsymbol{\Sigma}_Y$, denoted by $\boldsymbol{\Sigma}_1$, is of the form (3.3) with $v_i = 2$ for $i = 1, \ldots, 19$, $\phi_{i,i-1} = 0.5$ and $\phi_{i,j} = 0$ for $2 \leq j < i \leq 19$, $j \neq i-1$. This is an AR(1) structure that Huang et al. (2006) adopted in their simulation study. The second $\boldsymbol{\Sigma}_Y$, denoted by $\boldsymbol{\Sigma}_2$, is the covariance matrix implied by the structural equation model in Figure 2. Therefore, $\boldsymbol{\Sigma}_2$ mimics the covariance structure of the behavioral model observed in the survey data. To understand the effect of the distribution of $\boldsymbol{Y}$ on the performance of the Bayesian estimator, we set $\mathbf{Y}$ to follow two Gaussian copulas whose correlation parameters are specified by the standardized version of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$, and (i) normal margins and (ii) Poisson margins, with their marginal variances equal to the diagonal elements of $\boldsymbol{\Sigma}_j$. Hence, four experimental settings, $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ combined with normal and Poisson margins, are produced. We set $\mathbf{U}$ as a p-dimensional random vector independent of $\mathbf{Y}$, following a normal distribution with mean $\mathbf{0}$ and a diagonal covariance matrix having all entries equal to 2. We accompany the first p elements of $\mathbf{Y}$ by $\mathbf{U}$, and generate $\{Z_{k,i}\}_{i=1,\ldots,p}$ through the randomization procedure in (2.1). The remaining q elements of $\mathbf{Y}$ are defined as $\mathbf{D}$. Two samples with different ω_k are generated. Their sample sizes, n_1 and n_2, are both set to 200, so that the sample sizes here are comparable with those in Section 5. Moreover, we set $p = 7$, $q = 12$, $\omega_1 = 1/3$, and $\omega_2 = 2/3$ based on our settings in Section 5. We generate one hundred datasets under each experimental setting and implement the MCMC simulation algorithm described in Sections 3.3 and 4.1. We use Ando (2011)'s information criterion to set the shrinkage parameter λ to be 10.

We denote the MCMC estimate of the posterior mean of $\mathbf{\Sigma}_Y$ by $\widehat{\mathbf{\Sigma}}_Y$. To evaluate the accuracy of $\widehat{\mathbf{\Sigma}}_Y$, we consider two loss functions for covariance matrix estimators, known as entropy and quadratic losses (Muirhead, 2005, Chapter 4), which are defined by $\Delta_1(\mathbf{\Sigma}_Y, \widehat{\mathbf{\Sigma}}_Y) = \mathrm{tr}(\mathbf{\Sigma}_Y^{-1}\widehat{\mathbf{\Sigma}}_Y) - \log|\mathbf{\Sigma}_Y^{-1}\widehat{\mathbf{\Sigma}}_Y| - (p+q)$ and $\Delta_2(\mathbf{\Sigma}_Y, \widehat{\mathbf{\Sigma}}_Y) = \mathrm{tr}\{(\mathbf{\Sigma}_Y^{-1}\widehat{\mathbf{\Sigma}}_Y - \mathbf{I})^2\}$, respectively. The entropy and quadratic risks, defined as $R_j(\mathbf{\Sigma}_Y, \widehat{\mathbf{\Sigma}}_Y) = E(\Delta_j(\mathbf{\Sigma}_Y, \widehat{\mathbf{\Sigma}}_Y))$ for $j = 1, 2$, respectively, can be estimated by the mean loss of the replicated datasets. We use $R_j(\mathbf{\Sigma}_Y, \widehat{\mathbf{\Sigma}}_Y)$ to summarize the expected deviation of $\widehat{\mathbf{\Sigma}}_Y$ from $\mathbf{\Sigma}_Y$. We then compare the estimation performance of $\widehat{\mathbf{\Sigma}}_Y$ with the method-of-moments estimator (Kwan et al., 2010), denoted by $\widetilde{\mathbf{\Sigma}}_Y$, which is given by (2.2) but with the theoretical moments replaced by their empirical analogue. We also replicate datasets to evaluate $R_j(\mathbf{\Sigma}_Y, \widetilde{\mathbf{\Sigma}}_Y)$, $j = 1, 2$. Note that $\widetilde{\mathbf{\Sigma}}_Y$ may not be positive definite (P.D.), especially when the sample size is small. In those cases, we repeatedly replicate datasets and discard the datasets in which $\widetilde{\mathbf{\Sigma}}_Y$ is not P.D. until we can retain 100 datasets for comparing $\widehat{\mathbf{\Sigma}}_Y$ and $\widetilde{\mathbf{\Sigma}}_Y$. We obtain the relative efficiency (R.E.) of $\widehat{\mathbf{\Sigma}}_Y$ with respect to $\widetilde{\mathbf{\Sigma}}_Y$ under Δ_j, calculated by $R_j(\mathbf{\Sigma}_Y, \widetilde{\mathbf{\Sigma}}_Y)/R_j(\mathbf{\Sigma}_Y, \widehat{\mathbf{\Sigma}}_Y)$, to compare estimation performance. We also calculate the proportion of datasets with P.D. $\widetilde{\mathbf{\Sigma}}_Y$.

We set the initial values of $\{\phi_{i,j}\}$ and $\{\log v_i\}$ to 0 for all i and j, and $\mathbf{\Omega}_{t,k}$, the covariance of the proposal kernel for the kth sampling block of $\boldsymbol{\theta}$ at iteration t, to $c_{i(t,k)}\mathbf{I}$, where $i(t,k) = i$ if the kth sampling block is a subset of the block $(\phi_{i,1}, \ldots, \phi_{i,i-1}, \log v_i)$ and c_i is a positive constant. In other words, the $\phi_{i,j}$'s and $\log v_i$ in the same block of $(\phi_{i,1}, \ldots, \phi_{i,i-1}, \log v_i)$ share the same scaling constant c_i. We tune c_i to maintain the acceptance rates of the Markov chain at 25-50%. As the posterior variance of some $\log v_i$ is much different from that of the $\phi_{i,j}$'s under the same block, we may need to assign a specific scaling constant for $\log v_i$. Overall, we carry out 15,000 burn-in iterations followed by another 40,000 iterations for sample sizes of $n_1 = n_2 = 200$, which on average takes two hours in each replication.

TABLE 1
The R.E. and the proportion of datasets with P.D. $\widetilde{\mathbf{\Sigma}}_Y$ in the simulation study

	Normal		Poisson	
	$\mathbf{\Sigma}_1$	$\mathbf{\Sigma}_2$	$\mathbf{\Sigma}_1$	$\mathbf{\Sigma}_2$
R.E. under Δ_1	2.05	1.72	1.91	1.63
R.E. under Δ_2	1.40	1.46	1.38	1.48
Proportion of P.D. $\widetilde{\mathbf{\Sigma}}_Y$	36.5%	4.6%	40.5%	5.4%

4.2. *Evaluating the Bayesian estimator using relative efficiency.* We conduct a simulation study to evaluate the estimation performance of our Bayesian RRT method under the four data-generating processes described in Section 4.1. Table 1 reports the R.E. of $\widehat{\mathbf{\Sigma}}_Y$ with respect to $\widetilde{\mathbf{\Sigma}}_Y$ under Δ_j, $j = 1, 2$, and the P.D. proportion of $\widetilde{\mathbf{\Sigma}}_Y$ under each of the preceding four experimental settings. For the dependence structure given by $\mathbf{\Sigma}_1$, $\widehat{\mathbf{\Sigma}}_Y$ is 105%/91% and 40%/38% more efficient than $\widetilde{\mathbf{\Sigma}}_Y$ for $\mathbf{Y}$ following normal/Poisson margins under Δ_1 and Δ_2, respectively. For the dependence structure given by $\mathbf{\Sigma}_2$, $\widehat{\mathbf{\Sigma}}_Y$ is also 72%/63% and 46%/48% more efficient than $\widetilde{\mathbf{\Sigma}}_Y$ for $\mathbf{Y}$ following normal/Poisson margins under Δ_1 and Δ_2, respectively. These results show that our Bayesian RRT estimator is significantly more accurate than the method-of-moments estimator. Furthermore, the P.D. proportion of $\widetilde{\mathbf{\Sigma}}_Y$ is low in all of the experimental settings. When the dependence structure of $\mathbf{Y}$ is given by $\mathbf{\Sigma}_2$, the P.D. proportion is even as low as 5%, meaning that the method-of-moments estimator cannot provide valid estimates about 95% of the time. However, our Bayesian estimator does not suffer from the P.D. problem. The results demonstrate that the modified Cholesky decomposition greatly improves the regularity of the estimate of $\mathbf{\Sigma}_Y$, so that the positive definiteness can be achieved easily. Together with the shrinkage effect imposed by the Laplace prior of $\phi_{i,j}$, the estimation error can be reduced drastically.

4.3. *Effect of sample size on the performance.* In this section, we evaluate the performance of our Bayesian RRT method under different sample sizes. The settings here are the same as those in Section 4.1, except that we consider four different combinations of sample sizes with $n_1 = n_2 = n$ and $n = 200, 400, 600, 800$. For each combination of sample size, we generate 100 datasets using $\mathbf{\Sigma}_2$. We implement the MCMC algorithm described in Sections 3.3 and 4.1, and regard the posterior mean $\widehat{\mathbf{\Sigma}}_Y$ as the point estimator. We adopt the entropy and quadratic risks again to assess the estimation accuracy of $\widehat{\mathbf{\Sigma}}_Y$ and the benchmark method-of-moments estimator $\widetilde{\mathbf{\Sigma}}_Y$. Figure 1 presents the entropy and quadratic risks of $\widehat{\mathbf{\Sigma}}_Y$ and $\widetilde{\mathbf{\Sigma}}_Y$ under different combinations of sample size.

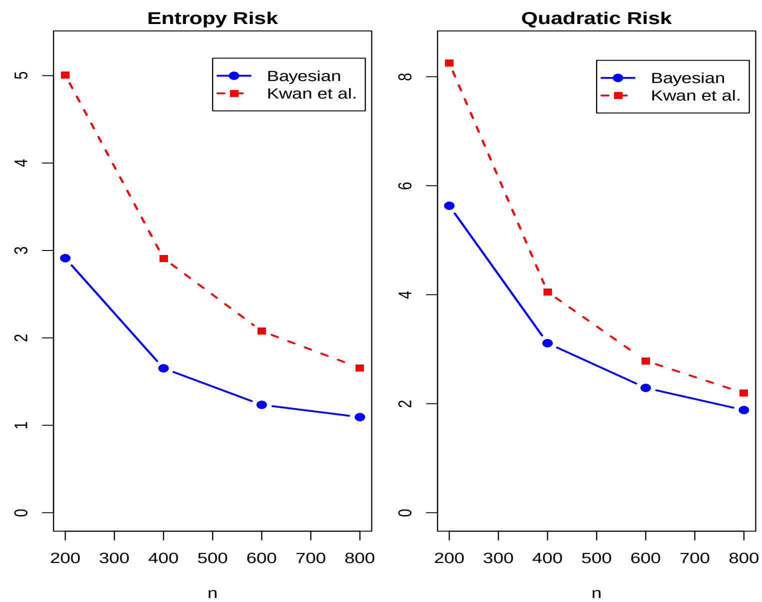

FIG 1. *Entropy and quadratic risks of the Bayesian RRT estimator and method-of-moments estimator of Kwan et al. (2010) under different sample sizes.*

When the sample size is small, such as $n = 200$, using $\widehat{\boldsymbol{\Sigma}}_Y$ instead of $\widetilde{\boldsymbol{\Sigma}}_Y$ reduces both entropy and quadratic risks by more than 30%. Nevertheless, the proportion of error reduction decreases with the sample size, which is not surprising as the curse of dimensionality is abated with sample size. Although the modified Cholesky decomposition and shrinkage effect mentioned in Section 4.2 no longer help much in terms of reducing Δ_1 and Δ_2 in the case of $n = 800$, the Bayesian RRT method still outperforms the method in Kwan et al. (2010) and avoids the P.D. problem. Based on the experiment, our Bayesian RRT method is more useful in cases with a small to moderate sample size, in which the curse of dimensionality is significant. If we want to keep the entropy or quadratic risk of estimation under a certain level, we can refer to Figure 1 to determine suitable sample sizes for both the Bayesian RRT method and the method-of-moments.

5. Empirical study.

5.1. *Motivation and model.* In this section, we apply the Bayesian RRT method developed in Section 3 to study IS resource misuse via structural equation modeling (SEM), a method extensively used in the social sciences literature to study the causal relationship between human behavior and its determinants. A brief introduction to SEM can be found in Kaplan (2009). We investigate the causal relationship between organizational commitment, punishment severity, attitude, and IS resource misuse.

Organizational commitment (Cohen, 1996; Jaros, 1997; Panaccio and Vandenberghe, 2009) and punishment severity (Peace et al., 2003) are two commonly investigated determinants of work-relevant activities in the social sciences literature. There are three forms of organizational commitment (Meyer et al., 1993), including (i) affective commitment (AFF), which refers to employees' emotional attachment to their organizations; (ii) continuance commitment (CON), which represents employees' awareness of the costs associated with leaving their organizations; and (iii) normative commitment (NOR), which denotes employees' feeling of obligation to remain in their organizations. Punishment severity (PUN) is defined as the fear of punishment from being caught misusing IS resources. We hypothesize that all of the aforementioned determinants affect one's attitude toward IS resources (ATT), which further determines the actual IS resource misuse behavior (ACT). We also hypothesize that the punishment severity has a direct effect on the actual IS resource misuse behavior. To describe the causal relationship mentioned previously, we construct a structural equation model with the form given in Figure 2. As the determinants and behavior are abstract, we may not be able to measure them directly using a single question. To address this problem, the standard SEM framework assumes that all of the determinants and behavior, including AFF, CON, NOR, PUN, ATT, and ACT, are latent variables represented by ovals in Figure 2. Each is measured by a set of questions with observable responses, represented by rectangles in Figure 2 (e.g., AFF1-3; ATT1-3).

5.2. *Survey design and data collection.* To fit the structural equation model in Figure 2, we first estimate the covariance matrix of the questions in the research model (represented by rectangles in Figure 2). Then, we input the estimated covariance matrix into an R function called `sem()` in R package `lavaan`, which is designed to fit a variety of latent variable models, to obtain the estimated path coefficients of the research model. The questions adopted in this study and their corresponding scale of answers are listed in Table 2. As the questions measuring attitude toward misuse and actual misuse behavior are sensitive, they are asked under the randomization procedure of the RRT. Hence, this study consists of $q = 12$ direct questions and $p = 7$ randomized questions. We collected the responses to the questions via an online survey.

We followed the advice of Kwan et al. (2010) to choose the two different probabilities of answering sensitive questions as $\omega_1 = 1/3$ and $\omega_2 = 2/3$ for

TABLE 2
Items Involved in the Study

Latent Variable	Item Description	Randomized	Scale#
Affective commitment (Meyer et al., 1993)			
AFF1	I would be very happy to spend the rest of my career with my organization.	No	A
AFF2	I feel a strong sense of belongingness to my organization.	No	A
AFF3	I feel like part of the family at my organization.	No	A
Continuance commitment (Meyer et al., 1993)			
CON1	Right now, staying with my organization is a matter of necessity.	No	A
CON2	It would be very hard for me to leave my organization right now, even if I wanted to.	No	A
CON3	Too much of my life would be disrupted if I decided to leave my organization now.	No	A
Normative commitment (Meyer et al., 1993)			
NOR1	Even if it were to my advantage, I do not feel it would be right to leave my organization now.	No	A
NOR2	I would feel guilty if I left my organization now.	No	A
NOR3	My organization deserves my loyalty.	No	A
Punishment severity (D'Arcy et al., 2009; Kwan et al., 2010)			
PUN1	If I were caught engaging in IS resource misuse, I would be severely reprimanded.	No	A
PUN2	If I were caught engaging in IS resource misuse, I would be severely punished.	No	A
PUN3	Even if I am caught engaging in IS resource misuse, I would not be subject to severe punishment. (Reverse)	No	A
Attitude toward the behavior* (Hsieh et al., 2008; Kwan et al., 2010)			
ATT1	Committing IS resource misuse is a bad/good idea. (I find saving money in time deposit is a bad/ good idea.)	Yes	B
ATT2	Committing IS resource misuse is a foolish/wise idea. (I find recording daily expenses in detail is a foolish/wise idea.)	Yes	B
ATT3	Committing IS resource misuse is harmful/ beneficial to the organization. (To me, taking vitamin pills every day is harmful/ beneficial.)	Yes	B
Actual misuse behavior* (Kwan et al., 2010; Chu and Chau, 2014)			
ACT1	Using untrusted network (e.g., the Internet) for data transmission at work. (Taking public transportation.)	Yes	C
ACT2	Installing untrusted applications for personal purposes at work. (Having dinner at home.)	Yes	C
ACT3	Running untrusted applications for personal purposes at work. (Going shopping.)	Yes	C
ACT4	Using instant messaging services at work without permission. (Singing Karaoke.)	Yes	C

* Statements in parenthesis are the unrelated questions paired with their respective sensitive questions.

A: 7-point Likert scale (Strongly agree - Strongly disagree), B: 7-point bipolar adjective scale, C: 7-point scale (Never - Very many times)

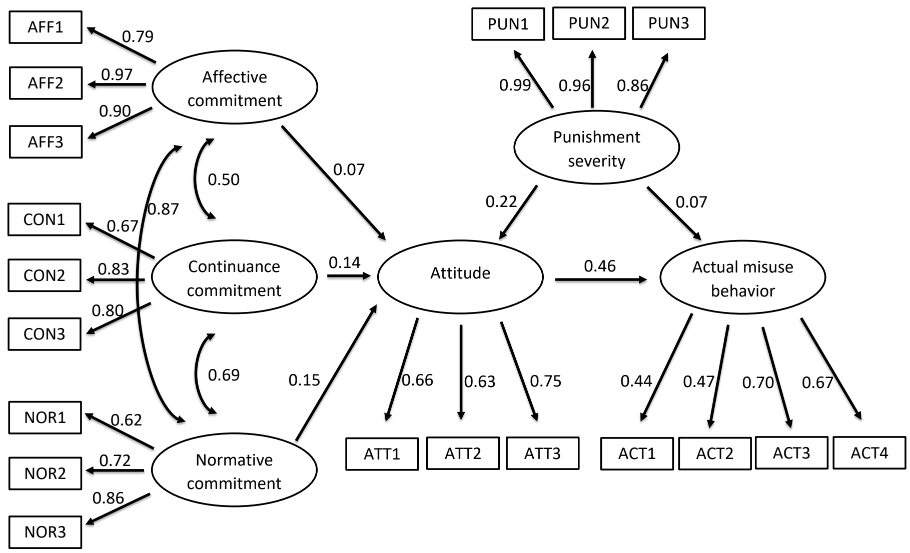

FIG 2. *Relations between latent variables under the structural equation model. All of the determinants and behavior, including AFF, CON, NOR, PUN, ATT, and ACT, are regarded as latent variables represented by ovals. Each is measured by a set of questions with observable responses, represented by rectangles (e.g., AFF1-3; ATT1-3).*

the two independent samples, i.e., Samples 1 and 2, respectively. To implement the randomization procedure in Sample 1, the respondents were asked to write down one of three numbers (e.g., 1, 2, 3) in secret on a piece of paper randomly and click a button on the screen to confirm the action. An example of the question format for a randomized item is shown in Figure 3. The online survey then generated a random number from the three numbers, and displayed the sensitive and the unrelated questions simultaneously. When the generated number matched with the secret number from the respondent (for $\omega_1 = 1/3$), the respondent answered the sensitive question. Otherwise, the respondent answered the unrelated question, which was innocuous and non-sensitive. The randomization procedure in Sample 2 was similar, except that respondents needed to answer the sensitive question in case the written number differed from, rather than coincided with, the random number (for $\omega_2 = 2/3$). With their own secret number, the respondents had more confidence in their privacy while enabling the randomization in the RRT, as nobody knew their secret number. Our procedure is preferable to the usual computer-assisted setting (Coutts and Jann, 2011), where the random number generated by the computer alone determines whether the respondents answer the sensitive question, because the respondents may not trust

Question

Instruction:

Please write down a number from **1, 2, 3** on a piece of paper randomly.

When completed, please click the button below and then click the "next" button to continue the survey.

O I've already written down a number from 1, 2, 3

A random number is generated by computer. It is **"2"**.

If this number matches the number you have just written down, please answer S; otherwise, please answer U.

RRT Question:

	Very Harmful **(1)**	**(2)**	**(3)**	Neutral **(4)**	**(5)**	**(6)**	Very Beneficial **(7)**
S. Committing IS resource misuse is (harmful/beneficial) to the organization. **U.** To me, taking vitamin pills every day is:	O	O	O	O	O	O	O

FIG 3. *Sample layout of an online page for the unrelated question design.*

the true randomness of the random number. We collected responses from a database of a marketing research firm. Finally, we used $n_1 = n_2 = 225$ completed responses in Samples 1 and 2.

5.3. *Results and implications.* After collecting the data, we centered the responses to the questions listed in Table 2 by the sample means so that the items were of mean zero. We then applied the MCMC method introduced in Section 3, with the posterior mean of $\mathbf{\Sigma}_Y$, i.e., the covariance matrix of the responses in the research model in the context of this section, considered as the point estimate of $\mathbf{\Sigma}_Y$. We inputted the point estimate of $\mathbf{\Sigma}_Y$ into the R function `sem()` mentioned previously, and the `sem()` returned a solution of the path coefficients of the research model. We divided the MCMC simulation into three stages. In the first stage, we set $\Omega_{t,k} = c_{i(t,k)}\mathbf{I}$, where $c_{i(t,k)}$ is a positive constant as defined in Section 4.1. We generated 10,000 burn-in observations followed by 10,000 observations. In the second stage, we set $\mathbf{\Omega}_{t,k}$ to be the scaled sample covariance matrix of the last 10,000 observations from the first stage of the MCMC simulation. Then, we generated another 10,000 burn-in observations followed by 10,000 observations. The third stage was similar to the second stage, in that we set $\mathbf{\Omega}_{t,k}$ to the scaled sample covariance matrix of the last 10,000 observations from the second stage, and generated 10,000 burn-in observations followed by 50,000 observations. The inference was solely based on the last 50,000 observations generated in the third stage. This adaptation method was applied in So et al. (2005). We set the shrinkage intensity parameter λ to 12.0, following

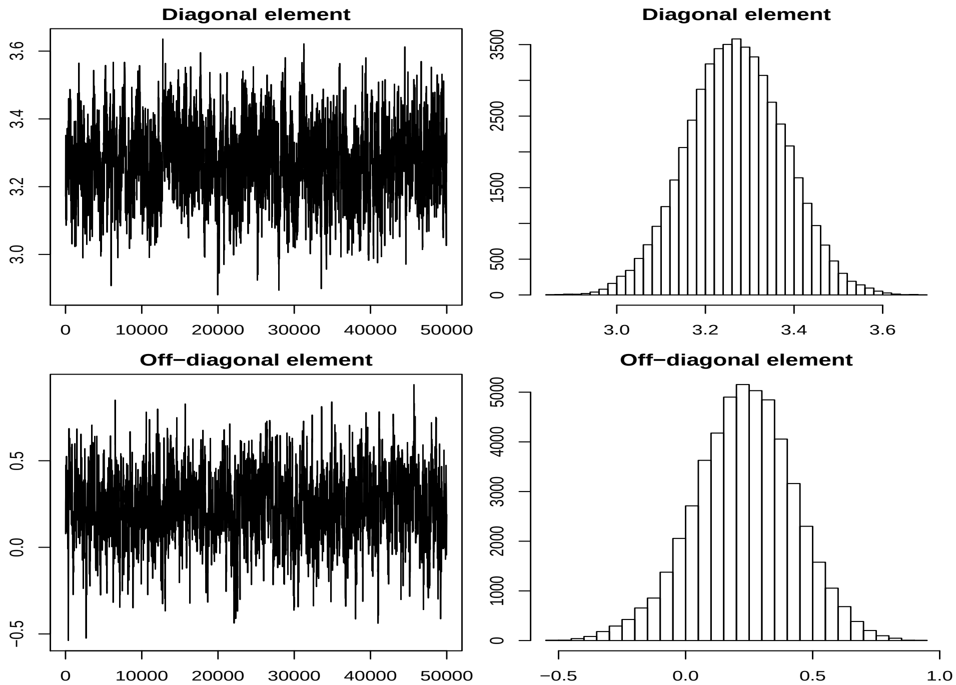

FIG 4. *Trace plots and histograms of the MCMC sample of some elements of $\mathbf{\Sigma}_Y$ in the empirical study.*

Ando (2011)'s information criterion.

Figure 4 shows the trace plots and histograms of the MCMC sample of a randomly selected diagonal element and a randomly selected off-diagonal element of $\mathbf{\Sigma}_Y$. From Figure 4, we can observe that the generated Markov chain reaches a stationary state after burn-in. Therefore, the stationary distribution of the chain provides a good approximation to the quasi-posterior distribution $\widehat{\pi}(\boldsymbol{\theta} \mid \mathbf{z}, \mathbf{d})$ of $\mathbf{\Sigma}_Y$. In the histograms, the quasi-posterior distribution of the elements of $\mathbf{\Sigma}_Y$ is in a bell shape, showing that the shape of the posterior distribution is proper even though the true likelihood is replaced by a quasi-likelihood. Table 3 presents the point estimate (i.e., the posterior mean estimated from the MCMC sampling) of the correlation matrix of the sensitive items (i.e., ATT1-3 and ACT1-4), whose responses are collected by the RRT. We also present the estimate obtained from the method-of-moments of Kwan et al. (2010) in the table. The correlations within the block of variables comprising ATT1-3 are high, with the magnitudes given by 0.45, 0.51, and 0.47. The correlations within the block of variables comprising ACT1-4 are significant, with all estimates greater than

Table 3
Correlation matrix estimate of the randomized items. The numbers in the parentheses are estimates from the method-of-moments

	ATT1	ATT2	ATT3	ACT1	ACT2	ACT3	ACT4
ATT1							
ATT2	0.45						
	(0.51)						
ATT3	0.51	0.47					
	(0.55)	(0.53)					
ACT1	0.08	0.36	0.25				
	(0.09)	(0.45)	(0.31)				
ACT2	0.01	0.17	0.18	0.34			
	(0.00)	(0.20)	(0.20)	(0.35)			
ACT3	0.24	-0.04	0.25	0.25	0.34		
	(0.24)	(-0.15)	(0.28)	(0.30)	(0.44)		
ACT4	0.04	0.15	0.39	0.26	0.26	0.50	
	(-0.01)	(0.13)	(0.44)	(0.32)	(0.29)	(0.53)	

0.2. In addition, the correlations of ATT1 with ACT3, ATT2 with ACT1, ATT3 with ACT3, and ATT3 with ACT4 are equal to 0.24, 0.36, 0.25, and 0.39, respectively, indicating that the latent factor characterized by ATT1-3 may be positively correlated with that characterized by ACT1-4. The estimate of the covariance matrix from the method-of-moments is different from that of our Bayesian method. Furthermore, the estimate from the method-of-moments is not positive definite, and hence we cannot proceed further to obtain the estimates of the path coefficients of the structural equation model. In short, our Bayesian method can simultaneously return a reasonable estimate of the covariance matrix and keep the estimate of the covariance matrix positive definite.

Figure 2 gives the estimated standardized path coefficients of the research model from `sem()`. The path coefficients from continuance commitment to attitude and from normative commitment to attitude are positive. Therefore, continuance and normative commitment have direct positive effects on the employees' attitudes toward misuse behavior. Affective commitment does not help much in weakening attitude toward misuse behavior, as observed from the corresponding near-to-zero path coefficient. Punishment severity affects employees' attitude toward misuse behavior directly, however, its effect on actual misuse behavior is very weak. As predicted, attitude is positively associated with actual misuse behavior. The relationships shown in Figure 2 provides insight into how the three forms of organizational commitment, and the punishment severity affect employees' engagement in IS resource misuse behavior. In fact, we observe from Table 3 that ATT1-3 and ACT1-4 are very much correlated, explaining why the attitude in Figure 2 has direct effect on the actual misuse behavior (path coefficient = 0.46). In short, we provide evidence using the structural equation model with sensitive responses that employees' attitude toward IS resource misuse, driven by continuance commitment (path coefficient = 0.14), normative commitment (path coefficient = 0.15) and punishment severity (path coefficient = 0.22), is a main determinant of the actual misuse behavior.

6. Conclusion and discussion.

Our study of IS resource misuse in the workplace reveals the importance of multivariate analysis on sensitive quantitative attributes. The RRT is a classical method for analysis of sensitive attributes, and the method-of-moments approach studied by Kwan et al. (2010) permits multivariate analysis of sensitive quantitative attributes. However, existing methods are unstable and probably return invalid covariance estimates. The problem is especially significant when the number of questions involved is large. This deficiency motivates us to introduce a new Bayesian RRT method.

In detail, we first reparameterize the covariance matrix of the attributes via a modified Cholesky decomposition, so that the intrinsic constraint of positive definiteness can be imposed naturally. We also impose a shrinkage effect on the estimate of the covariance matrix through the Laplace prior. Provided that the true covariance matrix is sparse, the shrinkage effect helps to decrease the estimation error. We demonstrate the advantage of using the shrinkage prior by the results in Section 4. A novelty of our Bayesian RRT method is its replacement of the true likelihood by a quasi-likelihood constructed from a set of moment equations. We also prove Bayesian consistency to theoretically justify our Bayesian RRT method using a quasi-likelihood. As the quasi-likelihood relies only on semi-parametric information instead of full parametric information, we can implement the Bayesian RRT method without specifying the joint distribution of the attributes.

We conduct a simulation to demonstrate that our Bayesian RRT method can smoothly be applied with the aid of a specially designed MCMC simulation algorithm, whose code is included in an R function called `BayeRRT()` for convenient application. We have also conducted other simulation studies to verify that our Bayesian RRT method outperforms Kwan et al. (2010) in dealing with multiple sensitive attributes. We apply our Bayesian RRT to investigate the causal relationships of three forms of organizational commitment, punishment severity, attitude toward IS resource misuse, and actual IS resource misuse behavior. With the aid of the SEM, we identify the relationships between different determinants of the actual IS resource misuse behavior.

We may estimate $\mathbf{\Sigma}_Y$ by projecting the method-of-moments estimator into the space of positive definite matrices. Nevertheless, when $\mathbf{Y}$ contains a moderately large number of questions, we need to optimize a high-dimensional objective function in estimation. This may result in an unreliable estimate that is trapped at either a local minimum or the boundary set of the parameter space. In contrast, our Bayesian RRT method returns a robust estimate of $\mathbf{\Sigma}_Y$ even though $\mathbf{Y}$ contains a moderately large or large number of questions. Hence, our method is preferable to projection methods when estimating the covariances among a large number of questions under the UQD settings. Although we focus on UQD in this paper, in principle, we can also use the Bayesian RRT method in other RRT designs. When adapting the method for another RRT design, the main thing is to rewrite Z_{ki} in terms of the indicator I_{ki} and the sensitive response S_i. In the mirrored question design of Warner (1965), we need only one sample and can write $Z_{1i} = I_{1i}S_i + (1 - I_{1i})(1 - S_i)$. For the forced response design (Blair et al., 2015), $Z_{1i} = I_{1i} + (1 - I_{1i} - I_{2i})S_i$, where the indicator variables $I_{1i} = I$(forced to say YES or 1) and $I_{2i} = I$(forced to say NO or 0) are defined. Then, we can derive all of the moment equations accordingly. The proposed Bayesian RRT method is not only able to deal with the RRT case, but also applicable to estimating covariance matrices with incomplete data information semiparametrically. It would also be interesting to investigate the Cramer-Rao lower bound (Singh and Sedory, 2011; Lee et al., 2016) for continuous responses in further research.

APPENDIX A: MOMENT EQUATIONS

Define the augmented unrelated question vector as $\widetilde{\mathbf{U}} = (\mathbf{U}^T, \mathbf{1}_q^T)^T$. Equation (2.1) can be expressed in

$$\mathbf{Z}_k = \begin{pmatrix} \mathbf{I}_k \\ \mathbf{1}_q \end{pmatrix} \odot \mathbf{Y} + \begin{pmatrix} \mathbf{I}_k^c \\ \mathbf{0}_q \end{pmatrix} \odot \widetilde{\mathbf{U}},$$

where $\mathbf{I}_k^c = \mathbf{1}_p - \mathbf{I}_k$ for $k = 1, 2$. Using the above relation, we have

$$\begin{aligned} \mathbf{Z}_k\mathbf{Z}_k^T = \quad & \left(\begin{pmatrix} \mathbf{I}_k \\ \mathbf{1}_q \end{pmatrix} (\mathbf{I}_k^T \mathbf{1}_q^T)\right) \odot \mathbf{Y}\mathbf{Y}^T + \left(\begin{pmatrix} \mathbf{I}_k^c \\ \mathbf{0}_q \end{pmatrix} (\mathbf{I}_k^{cT} \mathbf{0}_q^T)\right) \odot \widetilde{\mathbf{U}}\widetilde{\mathbf{U}}^T \\ & + \left(\begin{pmatrix} \mathbf{I}_k \\ \mathbf{1}_q \end{pmatrix} (\mathbf{I}_k^{cT} \mathbf{0}_q^T)\right) \odot \mathbf{Y}\widetilde{\mathbf{U}}^T + \left(\begin{pmatrix} \mathbf{I}_k^c \\ \mathbf{0}_q \end{pmatrix} (\mathbf{I}_k^T \mathbf{1}_q^T)\right) \odot \widetilde{\mathbf{U}}\mathbf{Y}^T. \end{aligned}$$

Taking expectation on both sides of the above equation and using the facts that $\mathbf{Y}$, $\mathbf{U}$, and $\mathbf{I}_k$ are independent, we can derive

$$E(\mathbf{Z}_k\mathbf{Z}_k^T) = \begin{pmatrix} \boldsymbol{\mathcal{P}}_k & \omega_{\mathbf{k}p\times q} \\ \omega_{\mathbf{k}q\times p} & \mathbf{1}_{q\times q} \end{pmatrix} \odot \boldsymbol{\Sigma}_Y + \begin{pmatrix} \boldsymbol{\mathcal{P}}_k^c \odot \boldsymbol{\Sigma}_U & \mathbf{0}_{p\times q} \\ \mathbf{0}_{q\times p} & \mathbf{0}_{q\times q} \end{pmatrix},$$

where $\omega_{\mathbf{k}p\times q}$ is a $p \times q$ matrix with all elements given by ω_k. Taking the operation $\mathbf{C}_k \odot \cdot$ on both sides of $E(\mathbf{Z}_k\mathbf{Z}_k^T)$ for $k = 1, 2$ and summing the resulting equations, we have $\mathbf{C}_1 \odot E(\mathbf{Z}_1\mathbf{Z}_1^T) + \mathbf{C}_2 \odot E(\mathbf{Z}_2\mathbf{Z}_2^T) = \mathbf{C}_0 \odot \boldsymbol{\Sigma}_Y$. As such, $\boldsymbol{\Sigma}_Y = \mathbf{C}_{1*} \odot E(\mathbf{Z}_1\mathbf{Z}_1^T) + \mathbf{C}_{2*} \odot E(\mathbf{Z}_2\mathbf{Z}_2^T)$.

APPENDIX B: PROOF OF THEOREM 3.1

Applying Taylor series expansion on $\mathbf{G}_n^{\mathcal{W}}(\boldsymbol{\theta})$ about $\widehat{\boldsymbol{\theta}}_n$ and given condition 1, we have $\mathbf{G}_n^{\mathcal{W}}(\boldsymbol{\theta}) = \dot{\mathbf{G}}_n^{\mathcal{W}}(\boldsymbol{\theta}_n^*)(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_n)(1 + o_p(1))$, where $\boldsymbol{\theta}_n^*$ lies between $\boldsymbol{\theta}$ and $\widehat{\boldsymbol{\theta}}_n$. In addition, by conditions 1, 3, and 4 in Theorem 1,

$$\dot{\mathbf{G}}_n^{\mathcal{W}}(\boldsymbol{\theta}_n^*) = \dot{\mathbf{G}}_n^{\mathcal{W}}(\boldsymbol{\theta}_0)\{\mathbf{I} + (\dot{\mathbf{G}}_n^{\mathcal{W}}(\boldsymbol{\theta}_0))^{-1}(\dot{\mathbf{G}}_n^{\mathcal{W}}(\boldsymbol{\theta}_n^*) - \dot{\mathbf{G}}_n^{\mathcal{W}}(\boldsymbol{\theta}_0))\} = \mathbf{H}_n(1 + o_p(1)).$$

Therefore, the posterior distribution of $\boldsymbol{\theta}$ can be expressed as

$$\begin{aligned} \widehat{\pi}(\boldsymbol{\theta} \mid \mathbf{z}, \mathbf{s}) &\propto \pi(\boldsymbol{\theta}) \exp\left\{-\frac{1}{2}(\mathbf{G}_n^{\mathcal{W}}(\boldsymbol{\theta}))^T \widehat{\mathbf{V}}_n^{-1} \mathbf{G}_n^{\mathcal{W}}(\boldsymbol{\theta})\right\} \\ &= \pi(\boldsymbol{\theta}) \exp\left\{-\frac{1}{2}(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_n)^T \mathbf{H}_n^T \mathbf{V}_n^{-1} \mathbf{H}_n (\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_n)(1 + o_p(1))\right\}. \end{aligned}$$

Let $\boldsymbol{\eta} = \boldsymbol{\Sigma}_n^{-1/2}(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}_n)$, where $\boldsymbol{\Sigma}_n = (\mathbf{H}_n^T \mathbf{V}_n^{-1} \mathbf{H}_n)^{-1}$. Denote the posterior distribution of $\boldsymbol{\eta}$ by $\widehat{\pi}(\boldsymbol{\eta} \mid \mathbf{z}, \mathbf{s})$. Then,

$$\widehat{\pi}(\boldsymbol{\eta} \mid \mathbf{z}, \mathbf{s}) \propto \pi(\widehat{\boldsymbol{\theta}}_n + \boldsymbol{\Sigma}_n^{1/2}\boldsymbol{\eta}) \exp\left\{-\frac{1}{2}\|\boldsymbol{\eta}\|^2(1 + o_p(1))\right\}. \tag{B.1}$$

Conditions 1 and 5 imply $\widehat{\boldsymbol{\theta}}_n + \boldsymbol{\Sigma}_n^{1/2}\boldsymbol{\eta} \xrightarrow{p} \boldsymbol{\theta}_0$, and hence $\pi(\widehat{\boldsymbol{\theta}}_n + \boldsymbol{\Sigma}_n^{1/2}\boldsymbol{\eta}) \xrightarrow{p} \pi(\boldsymbol{\theta}_0)$ by condition 6 and the continuous mapping theorem. Together with the fact that $\pi(\boldsymbol{\theta}_0) > 0$, we have

$$\widehat{\pi}(\boldsymbol{\eta} \mid \mathbf{z}, \mathbf{s}) \propto \pi(\boldsymbol{\theta}_0) \exp\left(-\frac{1}{2}\|\boldsymbol{\eta}\|^2\right)(1 + o_p(1)) \propto \exp\left(-\frac{1}{2}\|\boldsymbol{\eta}\|^2\right)(1 + o_p(1)).$$

The result follows.

REFERENCES

ANDO, T. (2011). Predictive Bayesian model selection. *American Journal of Mathematical and Management Sciences* **31** 13–38.

BLAIR, G. and IMAI, K. (2012). Statistical analysis of list experiments. *Political Analysis* **20** 47–77.

BLAIR, G., IMAI, K. and ZHOU, Y.-Y. (2015). Design and analysis of the randomized response technique. *Journal of the American Statistical Association* **110** 1304–1319.

BLAIR, G. and ZHOU, Y.-Y. (2016). Bayesian Randomized Response Regression. *The rr R package* [Online].
Available at https://github.com/SensitiveQuestions/rr/blob/master/R/rrBayes.R.

CHAUDHURI, A. (2011). *Randomized Response and Indirect Questioning Techniques in Surveys.* Chapman & Hall/CRC, Florida.

CHEN, C. C. and SINGH, S. (2011). Pseudo-Bayes and pseudo-empirical Bayes estimators in randomized response sampling. *Journal of Statistical Computation and Simulation* **81** 779–793.

CHRISTOFIDES, T. C. (2005). Randomized response technique for two sensitive characteristics at the same time. *Metrika* **62** 53–63.

COUTTS, E. and JANN, B. (2011). Sensitive questions in online surveys: Experimental results for the randomized response technique (RRT) and the unmatched count technique (UCT). *Sociological Methods & Research* **40** 169–193.

CHU, A. M. Y. and CHAU, P. Y. K. (2014). Development and validation of instruments of information security deviant behavior. *Decision Support Systems* **66** 93–101.

CHU, A. M. Y., CHAU, P. Y. K. and SO, M. K. P. (2015). Developing a typological theory using a quantitative approach: A case of information security deviant behavior. *Communications of the Association for Information Systems* **37** 510–535.

COHEN, A. (1996). On the discriminant validity of the Meyer and Allen measure of organizational commitment: How does it fit with the work commitment construct? *Educational and Psychological Measurement* **56** 494–503.

CRUYFF, M. J. L. F., VAN DEN HOUT, A. and VAN DER HEIJDEN, P. G. M. (2008). The analysis of randomized response sum score variables. *Journal of the Royal Statistical Society: Series B* **70** 21–30.

D'ARCY, J. and DEVARAJ, S. (2012). Employee misuse of information technology resources: Testing a contemporary deterrence model. *Decision Sciences* **43** 1091–1124.

D'ARCY, J., HOVAV, A. and GALLETTA, D. (2009). User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach. *Information Systems Research* **20** 79–98.

FOX, J. A. and TRACY, P. E. (1984). Measuring associations with randomized response. *Social Science Research* **13** 188–197.

FOX, J. A. and TRACY, P. E. (1986). *Randomized Response: A Method for Sensitive Surveys.* Sage Publications, California.

GJESTVANG, C. R. and SINGH, S. (2006). A new randomized response model. *Journal of the Royal Statistical Society: Series B* **68** 523–530.

GREENBERG, B. G., ABUL-ELA, A.-L. A., SIMMONS, W. R. and HORVITZ, D. G. (1969). The unrelated question randomized response model: Theoretical framework. *Journal of the American Statistical Association* **64** 520–539.

GREENBERG, B. G., KUEBLER, R. R., ABERNATHY, J. R. and HORVITZ, D. G. (1971). Application of the randomized response technique in obtaining quantitative data. *Journal of the American Statistical Association* **66** 243–250.

HANSEN, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* **50** 1029–1054.

Höglinger, M., Jann, B. and Diekmann, A. (2016). Sensitive questions in online surveys: An experimental evaluation of different implementations of the randomized response technique and the crosswise model *Survey Research Methods* **10** 171–187.

Horvitz, D. G., Shah, B. V. and Simmons, W. R. (1967). The unrelated question randomized response model. *Proceedings of Social Statistics Section, American Statistical Association* 65–72.

Hsieh, J. J. P.-A., Rai, A. and Keil, M. (2008). Understanding digital inequality: Comparing continued use behavioral models of the socio-economically advantaged and disadvantaged. *MIS Quarterly* **32** 97–126.

Huang, J. Z., Liu, N., Pourahmadi, M. and Liu, L. (2006). Covariance matrix selection and estimation via penalised normal likelihood. *Biometrika* **93** 85–98.

Imai, K. (2011). Multivariate regression analysis for the item count technique. *Journal of the American Statistical Association* **106** 407–416.

Imai, K., Park, B. and Greene, K. F. (2015). Using the predicted responses from list experiments as explanatory variables in regression models. *Political Analysis* **23** 180–196.

Jaros, S. J. (1997). An assessment of Meyer and Allen's (1991) three-component model of organizational commitment and turnover intentions. *Journal of Vocational Behavior* **51** 319–337.

Jayaraj, A., Odumade, O. and Singh, S. (2014). A New Quasi-Empirical Bayes Estimate in Randomized Response Technique. *JSM 2014*.

Kaplan, D. (2009). *Structural Equation Modeling*, 2nd ed. SAGE Publications, Inc.

Kuk, A. Y. C. (1990). Asking sensitive questions indirectly. *Biometrika* **77** 436–438.

Kwan, S. S. K., So, M. K. P. and Tam, K. Y. (2010). Research note - Applying the randomized response technique to elicit truthful responses to sensitive questions in IS research: The case of software piracy behavior. *Information Systems Research* **21** 941–959.

Lee, C. S., Sedory, S. A., and Singh, S. (2016). Cramer-Rao lower bounds of variance for estimating two proportions and their overlap by using two-decks of cards. *Handbook of Statistics* **34** 353–385.

Lin, C.-P. and Ding, C. G. (2003). Modeling information ethics: The joint moderating role of locus of control and job insecurity. *Journal of Business Ethics* **48** 335–346.

Liu, J. S., Liang, F. and Wong, W. H. (2000). The multiple-try method and local optimization in metropolis sampling. *Journal of the American Statistical Association* **95** 121–134.

Locander, W., Sudman, S. and Blackburn, N. (1976). An investigation of interview method, threat and response distortion. *Journal of the American Statistical Association* **71** 269–275.

Mangat, N. S. (1994). An improved randomized response strategy. *Journal of the Royal Statistical Society: Series B* **56** 93–95.

Mangat, N. S. and Singh, R. (1990). An alternative randomized response procedure. *Biometrika* **77** 439–442.

Meyer, J. P., Allen, N. J. and Smith, C. A. (1993). Commitment to organizations and occupations: Extension and test of a three-component conceptualization. *Journal of Applied Psychology* **78** 538–551.

Minsky-Kelly, D., Hamberger, L. K. and Pape, D. A. and Wolff, M. (2005). We've had training, now what?: Qualitative analysis of barriers to domestic violence screening and referral in a health care setting. *Journal of Interpersonal Violence* **20** 1288–1309.

Muirhead, R. J. (2005). *Aspects of Multivariate Statistical Theory.* Wiley-Interscience, New Jersey.

Nuno, A. and St John, F.A.V. (2015). How to ask sensitive questions in conservation: A review of specialized questioning techniques. *Biological Conservation* **189** 5–15.

Panaccio, A. and Vandenberghe, C. (2009). Perceived organizational support, organizational commitment and psychological well-being: A longitudinal study. *Journal of Vocational Behavior* **75** 224–236.

Park, T. and Casella, G. (2008). The Bayesian lasso. *Journal of the American Statistical Association* **103** 681–686.

Peace, A. G., Galletta, D. F. and Thong, J. Y. L. (2003). Software piracy in the workplace: A model and empirical test. *Journal of Management Information Systems* **20** 153–177.

Pollock, K. H. and Bek, Y. (1976). A comparison of three randomized response models for quantitative data. *Journal of the American Statistical Association* **71** 884–886.

Raghavarao, D. and Federer, W. T. (1979). Block total response as an alternative to the randomized response method in surveys. *Journal of the Royal Statistical Society: Series B* **41** 40–45.

Rosenfeld, B., Imai, K. and Shapiro, J. N. (2016). An empirical validation study of popular survey methodologies for sensitive questions. *American Journal of Political Science* **60** 783–802.

Shao, J. (2003). *Mathematical Statistics*, 2nd ed. Springer, New York.

Shephard, N. and Pitt, M. K. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika* **84** 653–667.

Singh, S. (2003). *Advanced Sampling Theory with Applications: How Michael 'Selected' Amy, Volume I.* Springer, New York.

Singh, S. and Sedory, S. A. (2011). Cramer-Rao lower bound of vvariance in randomized response sampling. *Sociological Methods & Research* **40** 536–546.

So, M. K. P., Chen, C. W. S. and Chen, M.-T. (2005). A Bayesian threshold nonlinearity test for financial time series. *Journal of Forecasting* **24** 61–75.

Tamhane, A. C. (1981). Randomized response techniques for multiple sensitive attributes. *Journal of the American Statistical Association* **76** 916–923.

Tan, M. T., Tian, G.-L. and Tang, M.-L. (2009). Sample surveys with sensitive questions: A nonrandomized response approach. *The American Statistician* **63** 9–16.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association* **60** 63–69.

Yin, G. (2009). Bayesian generalized method of moments. *Bayesian Analysis* **4** 191–208.

Division of Environment
The Hong Kong University of
Science and Technology
Clear Water Bay, Kowloon
Hong Kong E-mail: swchungaa@connect.ust.hk

Department of Mathematics and Statistics
Hang Seng Management College
Hang Shin Link, Siu Lek Yuen
Shatin, N.T., Hong Kong
E-mail: amandachu@hsmc.edu.hk

Department of Information Systems,
Business Statistics and Operations Management
The Hong Kong University of
Science and Technology
Clear Water Bay, Kowloon
Hong Kong E-mail: immkpso@ust.hk