Adjusting nonresponse bias in small area estimation without covariates via a Bayesian spatial model

Xiaoming Gao*, Chong He, and Dongchu Sun

Sometimes a survey sample is drawn from a large area even if the estimate of interest is at a smaller subdomain level. This strategy, however necessary, may cause small sample problems. The estimation problem is further complicated by survey nonresponse. We build a Bayesian hierarchical spatial model that takes into account both small sample size and nonresponse. This Bayesian model gives the estimates of marginal satisfaction rates at subdomains even when there is no covariate available via modeling the phase-specific response rates and conditional satisfaction rates given response status at subdomains. This method is illustrated using data from the 2001 Missouri Deer Hunter Attitude Survey. Satisfaction, in this survey, refers to whether respondents were satisfied with the Missouri Department of Conservation's deer management program. The estimated satisfaction rates are lower after adjusting for nonresponse bias compared to the satisfaction rates based only on responses.

Keywords and phrases: Nonresponse bias, Small area estimation, Attitude survey, Bayesian hierarchical model, Autoregressive models, Spatial random effect.

1. INTRODUCTION

Survey response rates have been declining for several decades at the very least (de Leeuw and de Heer, 2002). Estimates based on the response alone lead to biases. Lessler and Kalsbeek (1992) outlined partial solutions through questionnaire design and implementation. One common practice when conducting mail surveys is to resend the survey to nonrespondents multiple times in an effort to increase response rates (Dillman, 2000).

When there are multiple response phases in a survey, one of the most important questions is how the late responses are utilized in the data analysis. Lessler and Kalsbeek (1992) reviewed many frequency adjustment procedures in formulating and producing survey estimates. Bayesian approaches for dealing with nonresponse have been examined, for example, by Basu and de Braganca Pereira (1982), Albert and Gupta (1985), and Forster and Smith (1998). However, the work

*Corresponding author.

presented in these articles require large sample sizes, which is not always available. Today sample surveys are widely used in practice to provide estimates not only for the total population of interest but also for a variety of subpopulations (subdomains) (Rao, 2003). Sometimes sample sizes may be sufficient for inference of the total population, but still too small to obtain accurate estimates for subdomains. This is the small area estimation problem. Stasny (1991) modeled the success rate, response rate given failure, and response rate given success for each subdomain, and used an empirical Bayes procedure in finding Bayesian posterior distributions. A full Bayesian approach was extended and proposed in Nandram and Choi (2002b), who used Bayes factors for comparing ignorable and nonignorable models. Nandram and Choi (2002a) and Nandram et al. (2005) also used hierarchical Bayesian models for binary data or categorical data from small subdomains, but they considered neither multiple response phrases nor spatial dependency among small subdomains.

Making estimates for small subdomains with adequate levels of precision often requires the use of "indirect" estimators that borrow information from temporal or related geographic regions (Rao, 2003). For data collected over several geographic regions, such as counties in a state, a typical approach is to include spatial random effects in the first stage of a Bayesian hierarchical model. Conditional autoregressive (CAR) (Besag, 1974) or similar models are used to fit the spatial effects. Gao et al. (2014) used a bivariate Bayesian hierarchical lineal mixed model to estimate hunter's satisfaction rate and response rate. The estimated satisfaction and response rates were then aggregated at subdomain and statewide levels. Gao et al. (2014) explicitly included nonrespondents through auxiliary covariates, but these kind of covariates are not always available or hard to get in mail surveys. Oleson and He (2008) proposed a Bayesian hierarchical model to estimate hunting success rates given response phase and other phase-specific response rates. They used the simultaneous CAR model of Clayton and Kaldor (1987) to fit the spatial random effects in a generalized linear mixed model for phase specific success rates, but assumed phase-specific response rates were independently distributed and that there was no spatial dependency among small subdomains.

In this paper, we propose a bivariate Bayesian hierarchical spatial model to estimate satisfaction rates via model-

ing phase-specific response rates and conditional satisfaction rates given response statuses. Our approach is to utilize information from multiple response phases and neighboring small subdomains in adjusting nonresponse bias. The proposed model includes response phase as fixed effects and random spatial effects following a 2-fold CAR model (Woodard, 1999), which not only borrows strength from neighboring small subdomains, but also allows additional information sharing between response and satisfaction rates. This model does not included any other auxiliary covariates.

The data from the 2001 Missouri Deer Hunter Attitude Survey (MDHAS) is analyzed in this paper. The goal is to estimate 2001 firearms deer hunting permit buyers' satisfaction rates regarding the Missouri Department of Conservation (MDC) deer management program at both state and county levels in the presence of nonresponse. In this analysis, nonrespondents are those who did not return the 2001 MDHAS as well as those who returned the survey but did not answer the question regarding satisfaction with MDC's deer management program. Among respondents the satisfaction with MDC's deer management program is defined as "satisfied" or "unsatisfied", a binary outcome. The permit buyers' satisfaction rates will be estimated at their Missouri county of residence.

The rest of the paper is organized as follows. In Section 2, a brief description of the data is given and summarized at subdomain (county) levels. In Section 3, a bivariate Bayesian hierarchical spatial model is specified. Multiple response phases are fixed effects and county effects are spatial random effects. Prior distributions for the parameters are given. Bayesian computation via Gibbs sampling (Gelfand and Smith, 1990) is described in Section 4. The full conditional distributions are given in the Appendix. In Section 5 the 2001 MDHAS is analyzed. The article concludes with a discussion in Section 6.

2. DATA

One of the objectives of the MDC's deer management program is to maintain the deer population at a level that provides maximum hunting and viewing opportunities but minimizes problems associated with deer. Social surveys are conducted to gain a better understanding of residents and their relationship with wildlife (Filion, 1980). The 2001 MDHAS was designed to determine characteristics of 2001 firearms deer permit buyers along with their attitudes and perceptions concerning deer populations, hunting, and management in Missouri. MDC conducted a mail survey following Dillman (2000). The questionnaires were initially mailed to a simple random sample of 2001 firearms deer permit buyers on February 26, 2002, with two follow-up mailings on April 3 and May 16. Excluding the undeliverable questionnaires and non-Missouri residents, the total sample size was 9,040. About 32.3% of the hunters responded to the initial mailing. The 6,080 hunters who did not reply to the initial mailing received the first follow-up mailing, and 23% of those responded. The 4,634 hunters who did not reply to the initial

and the second mailings received the last follow-up mailing, and only 13% of those returned questionnaires. Overall, 55% responded to the 2001 MDHAS. Among those who returned the survey, 19 hunters did not answer the satisfaction question about MDC's deer management program. In general, the item nonresponse is different from the unit nonresponse. But we treat these 19 item nonresponses in the 2001 MD-HAS data the same as the unit nonresponse after a detailed examination. Nine of them were the second mailing respondents and are similar to the unit nonrespondents in the 2001 MDHAS case. The other 10 were the first mailing respondents. They did not answer several other opinion questions regarding deer hunting and deer management in the survey either. So, we consider them similar to the unit nonresponse in terms of satisfaction towards MDC's deer management program. Therefore these 19 are defined as nonrespondents in the analysis of satisfaction rate. This further reduced the response rate for the satisfaction question.

In the 2001 MDHAS, hunters were asked to rank their satisfaction on a scale from 1 to 5. A ranking of 1, 2, 3, 4, and 5 corresponded to excellent, good, fair, poor, and no opinion, respectively. Among the 4,968 respondents, 24 marked more than one ranking, while 198 had no opinion about deer management in Missouri. Those 222 respondents are excluded from the analysis. Therefore the response rate for the question regarding satisfaction of MDC's deer management program is 53.82%. Rankings of excellent or good are defined as satisfied with MDC's deer management program, while rankings of fair and poor are defined as not satisfied in this analysis. After dividing the data across three response phases and 114 counties, the number of respondents are as low as 1 or 0 for some counties. For example, Dade County and seven other counties have no data for the second and the third response phases. Cooper and three other counties have only one respondent for the second response phase. Caldwell and twelve other counties have only one respondent for the third response phase.

For each county i, let m_{i1} be the total number of questionnaire mailed in the first phase, n_{ij} ($j = 1, 2, 3$) be the number of respondents at response phase j, and y_{ij1} ($j = 1, 2, 3$) be the number of hunters among respondents who were satisfied with the MDC's deer management program at response phase j. Let n_{i4} be the number of nonrespondents and y_{i41} be the number of hunters who were satisfied with the MDC's deer management program among the nonrespondents. The relationship of these variables is better described by a two-way contingency table for the ith county in Table 1.

The marginal entries s_{i1} and s_{i2} are row sums for the number of satisfied and the number of unsatisfied hunters, respectively. Note that in Table 1, y_{i41} and y_{i42} are unobserved; thus s_{i1} and s_{i2} are also unobserved. The corresponding probabilities are presented in Table 2.

One way to obtain an estimate of satisfaction rate $p_i (= p_{i11} + p_{i21} + p_{i31} + p_{i41})$ is through conditional probabilities. We use the pattern-mixture model approach (Little and Ru-

Table 1. A two-way contingency table for observed and unobserved data

	Phase 1	Phase 2	Phase 3	Nonresponse	Total
Satisfied	y_{i11}	y_{i21}	y_{i31}	y_{i41}	s_{i1}
Unsatisfied	y_{i12}	y_{i22}	y_{i32}	y_{i42}	s_{i2}
Response	n_{i1}	n_{i2}	n_{i3}	n_{i4}	m_{i1}

Table 2. A two-way table of probabilities of observed and unobserved data

	Phase 1	Phase 2	Phase 3	Nonresponse	Total
Satisfied	p_{i11}	p_{i21}	p_{i31}	p_{i41}	p_i
Unsatisfied	p_{i12}	p_{i22}	p_{i32}	p_{i42}	$1 - p_i$
Response	q_{i1}	q_{i2}	q_{i3}	$1 - q_{i1} - q_{i2} - q_{i3}$	1

Table 3. A two-way table of conditional probabilities for observed and unobserved data

	Phase 1	Phase 2	Phase 3	Nonresponse	Total
Satisfied	$\pi_{i1}q_{i1}$	$\pi_{i2}q_{i2}$	$\pi_{i3}q_{i3}$	$\pi_{i4}(1 - \sum_{j=1}^{3} q_{ij})$	p_i
Unsatisfied	$(1-\pi_{i1})q_{i1}$	$(1-\pi_{i2})q_{i2}$	$(1-\pi_{i3})q_{i3}$	$(1-\pi_{i4})(1 - \sum_{j=1}^{3} q_{ij})$	$1 - p_i$
Response	q_{i1}	q_{i2}	q_{i3}	$1 - \sum_{j=1}^{3} q_{ij}$	1

bin, 2002) and factor the joint probability of response and satisfaction as well as the joint probability of nonresponse and satisfaction as

$$P(\text{response, satisfaction})$$
$$= P(\text{response})P(\text{satisfaction} \mid \text{response}),$$
$$P(\text{nonresponse, satisfaction})$$
$$= P(\text{nonresponse})P(\text{satisfaction} \mid \text{nonresponse}).$$

Define, for $i = 1, \ldots, I$

$$\pi_{ij} = P(\text{satisfaction} \mid \text{response to phase } j), \quad j = 1, 2, 3,$$
$$\pi_{i4} = P(\text{satisfaction} \mid \text{nonresponse}).$$
$$q_{ij} = P(\text{response to phase } j), \quad j = 1, 2, 3.$$

The relationship of conditional probabilities are shown in Table 3.

Our approach is first to estimate the conditional satisfaction rates π_{ij} for $j = 1, \ldots, 4$ and response rate q_{ij} for $j = 1, 2, 3$ and characterize their spatial dependencies. Once estimates for π_{ij} and q_{ij} are obtained, p_i can be estimated by using their relationship in Table 3.

3. HIERARCHICAL MODELS

3.1 Likelihood functions

For each subdomain i, let $\mathbf{y}_i = (y_{i11}, \ldots, y_{i41}, y_{i12}, \ldots, y_{i42})$, and $\mathbf{p}_i = (p_{i11}, \ldots, p_{i41}, p_{i12}, \ldots, p_{i42})$, $i = 1, \ldots, I$ (the total number of subdomain of interest), assume that

$$(\mathbf{y}_i \mid m_{i1}, \mathbf{p}_i) \stackrel{ind}{\sim}$$
$$Multinomial(m_{i1}, p_{i11}, p_{i21}, p_{i31}, p_{i41}, p_{i12}, p_{i22}, p_{i32}, p_{i42}).$$

The augmented likelihood function can be written as

$$(1) \quad [\mathbf{y}_i \mid m_{i1}, p_{i11}, p_{i21}, p_{i31}, p_{i41}, p_{i12}, p_{i22}, p_{i32}, p_{i42}]$$
$$= (m_{i1}!/\prod_{k=1}^{2}\prod_{j=1}^{4} y_{ijk}!) \prod_{k=1}^{2}\prod_{j=1}^{4} p_{ijk}^{y_{ijk}}$$
$$\propto \binom{n_{i4}}{y_{i41}} \Big(\prod_{j=1}^{4} \pi_{ij}^{y_{ij1}} (1-\pi_{ij})^{y_{ij2}} \Big)$$
$$\times \Big(\prod_{j=1}^{3} q_{ij}^{n_{ij}} (1 - \sum_{j=1}^{3} q_{ij})^{n_{i4}} \Big).$$

We reparameterize the response probabilities (q_{i1}, q_{i2}, q_{i3}) into discrete hazard rates or conditional probabilities to simplify the likelihood function. Define

$$(2) \quad h_{ij} = \frac{q_{ij}}{q_{ij} + \cdots + q_{i,J+1}}, \; j = 1, \ldots, J,$$

where $q_{i,J+1} = 1 - \sum_{i=1}^{J} q_{ij}$, $J = 3$. In the example of 2001 MDHAS, $h_{i1}(= q_{i1})$ is the response rate of phase 1, h_{i2} is the conditional probability of response at phase 2 given that the hunter didn't return the survey from phase 1, and h_{i3} is the conditional probability of response at phase 3 given that

the hunter didn't return the survey from the previous two phases. The likelihood function (1) now becomes a function of the conditional probability of satisfaction and the conditional probability of response, and is proportional to

$$(3) \qquad \propto \binom{n_{i4}}{y_{i41}} \Big(\prod_{j=1}^{4} \pi_{ij}^{y_{ij1}} (1 - \pi_{ij})^{n_{ij} - y_{ij1}} \Big)$$
$$\times \Big(\prod_{j=1}^{3} h_{ij}^{n_{ij}} (1 - h_{ij})^{m_{i1} - \sum_{l=1}^{j} n_{il}} \Big).$$

Again, the number of satisfied nonrespondents are unobserved. We treat y_{i41} as a latent variable.

3.2 Logistic linear mixed models

The logit link is the most common link for binary data. Probit and complementary log-log links are other frequently used links. Sometimes t-link is used as well. For a given mailing phase j, we compared the logit, probit, complementary log-log and t transformations for the observed satisfaction and response rates. The logit and probit transformed data fit normal much better than complementary log-log transformed data did. T transformed data fit normal well when its degree freedom is large which is essential the same as a probit transformation. Since our data are fairly symmetric, a skewed t-link (Chen et al., 1999) is not considered in this case. Final results from using logit and probit links are very similar for satisfaction, response rates at both subdomain and statewide levels. We choose logit link for easy interpretation. Therefore we propose the following linear mixed models to estimate π_{ij} and h_{ij},

$$(4) \qquad u_{ij} \equiv \log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = \theta_{1j} + Z_{1i} + e_{1ij},$$

$$(5) \qquad v_{ij} \equiv \log\left(\frac{h_{ij}}{1 - h_{ij}}\right) = \theta_{2j} + Z_{2i} + e_{2ij},$$

where θ_{1j} and θ_{2j} are the response phase effects $j = 1, 2, 3$, Z_{1i} and Z_{2i} are the spatial effects, $i = 1, \ldots, I$. Furthermore the θ_{1j} and θ_{2j} are treated as fixed effects with prior distributions defined later in this section, the Z_{1i} and Z_{2i} are treated as random effects with prior distribution defined in section 3.3, and the e_{1ij} and e_{2ij} are the phase-county random effects with

$$(6) \qquad e_{1ij} \stackrel{ind}{\sim} N(0, \delta_{1j}), \ j = 1, 2, 3,$$

$$(7) \qquad e_{2ij} \stackrel{ind}{\sim} N(0, \delta_{2j}), \ j = 1, 2, 3.$$

Note equation (4) defines π_{i1}, π_{i2} and π_{i3} only. The question of how to model π_{i4}, satisfaction rate given nonresponse is an important point of inquiry. It may be done by considering the possible linkage between response phase 2, 3 and nonresponse for a given data set or a survey problem. Filion (1980) suggested one approach was to use the replies to follow up mailings as a proxy for the nonrespondents. The means of logit of the empirical satisfaction rates are 1.4213,

1.0913, and 1.1818 for phase 1, 2, and 3 respondents, respectively. Note that the means for phase 2 and 3 are quite close. This indicates that nonrespondents may have a similar mean as the phase 3 respondents. To avoid the strong assumption that satisfaction rates of nonrespondents (or its logit) are the same as that of phase 3 respondents, we add a different phase-county random effect term in the equation (8) to model π_{i4}. The final data analysis results show that the phase-county random effect term for nonorespondents has larger variance, and this yields lower satisfaction rates. In Gao et al. (2014), the same survey was analyzed by using a different approach with individual covariates; their estimated statewide satisfaction rate of nonrespondents was 0.7100, which is very close to our estimate 0.7092 in this paper. This further indicates our model is reasonable for the data. Therefore, we have

$$(8) \qquad u_{i4} \equiv \log\left(\frac{\pi_{i4}}{1 - \pi_{i4}}\right) = \theta_{13} + Z_{1i} + e_{1i4},$$

where

$$(9) \qquad e_{1i4} \stackrel{ind}{\sim} N(0, \delta_{14}), \quad i = 1, \ldots, I.$$

In this way, $\pi_{i1}, \pi_{i2}, \pi_{i3}$, and π_{i4} have the same underlying spatial structure Z_{1i}. In fact, π_{i3} and π_{i4} are the same up to their respective phase-county random effects.

The usual way of modeling a probit or a logit model for small area estimation is through the insertion of one random term, a random area effect. But we include two random terms, one for random county effect and the other one for random phase-county interaction effect in equations (4), (5), and (8). An alternative model is to replace the phase-county random effect by relevant covariates if available.

In addition we assume that θ_{1j} and θ_{2j} are independent and with normal distributions

$$\theta_{1j} \stackrel{ind}{\sim} N(\mu_{1j}, \tau_{1j}),$$
$$\theta_{2j} \stackrel{ind}{\sim} N(\mu_{2j}, \tau_{2j}),$$

where $\mu_{1j}, \tau_{1j}, \mu_{2j}$, and $\tau_{2j}, j = 1, 2, 3$ are fixed constants.

3.3 Joint prior distribution of spatial effects

The conditional autoregressive (CAR) models (Besag, 1974) are popular for spatial data collected over irregular lattice structures such as counties in Missouri. However, CAR models can be appropriate in univariate cases. Recall in section 3.2, Z_{1i} and Z_{2i} are spatial effects in modeling conditional satisfaction rates given response status and phase specific response rates respectively. To model these two spatial parameters jointly, a natural choice is a 2-fold CAR model. We use the form provided in Woodard (1999) and the following specification is used.

Let $\mathbf{Z}_1 = (Z_{11}, \ldots, Z_{1I})^T, \mathbf{Z}_2 = (Z_{21}, \ldots, Z_{2I})^T$, and $\mathbf{Z} = (\mathbf{Z}_1^T, \mathbf{Z}_2^T)^T$. We say that $\mathbf{Z}$ follows a 2-fold CAR model if

$$(10) \qquad \Big(\mathbf{Z} \mid \delta_1, \delta_2, \rho_1, \rho_2, \rho_3 \Big) \sim N_{2I}(\mathbf{0}, \mathbf{\Sigma}),$$

where

$$(11) \quad \mathbf{\Sigma} = \begin{bmatrix} \delta_1 \mathbf{B}_1^{-1} & \rho_3\sqrt{\delta_1\delta_2}\mathbf{B}_1^{-\frac{1}{2}}\mathbf{B}_2^{-\frac{1}{2}} \\ \rho_3\sqrt{\delta_1\delta_2}\mathbf{B}_2^{-\frac{1}{2}}\mathbf{B}_1^{-\frac{1}{2}} & \delta_2\mathbf{B}_2^{-1} \end{bmatrix},$$

$$\mathbf{B}_k = \mathbf{I} - \rho_k\mathbf{C}, \quad k = 1, 2,$$

and $\mathbf{I}$ is a $I \times I$ identity matrix, $\mathbf{C} = (C_{lm})$ is a $I \times I$ symmetric adjacency matrix, $C_{lm} = 1$ if counties m and l share a common boundary, $C_{lm} = 0$ otherwise, including $C_{ll} = 0$. δ_1 and δ_2 are variance components. ρ_1 and ρ_2 are the spatial dependency parameters in $\mathbf{B}_1$ and $\mathbf{B}_2$, respectively. ρ_3 is a "bridging" or "linking" parameter between two spatial effects and is bounded between -1 and 1. Let $\lambda_1 \leq \cdots \leq \lambda_I$ be the eigenvalues of the matrix $\mathbf{C}$. $\mathbf{B}_k$ is positive definite if $\lambda_1^{-1} < \rho_k < \lambda_I^{-1}$ and $\lambda_I^{-1} > 0$, for $k = 1, 2$. For 2001 MDHAS, $I = 114$ and the numerical values of λ_1 and λ_I are -2.8931 and 5.6938, respectively (He and Sun, 2000). It means that the density of $\mathbf{Z}$ exists if ρ_1 and ρ_2 are in $(-0.3457, 0.1756)$.

Uniform prior distributions are assigned for spatial dependency parameters, ρ_1, ρ_2, and ρ_3, on their corresponding ranges as follows:

$$\begin{aligned} \rho_j &\sim U(\lambda_1^{-1}, \lambda_I^{-1}), \; j = 1, 2, \\ \rho_3 &\sim U(-1, 1). \end{aligned}$$

3.4 Prior distributions of variance components

In the model, there are a few variance components, $\delta_{1j}, \delta_{2j}, \delta_1$, and δ_2. δ_{1j} and δ_{2j} are for the residual effects of (6), (7), and (9). δ_1 and δ_2 are for the error terms in the variance structure of the random effects (11). These variance components are assumed to have inverse gamma distributions with prespecified hyperparameters,

$$\begin{aligned} \delta_{1j} &\overset{ind}{\sim} IG(a_{1j}, b_{1j}), \; j = 1, \ldots, 4, \\ \delta_{2j} &\overset{ind}{\sim} IG(a_{2j}, b_{2j}), \; j = 1, 2, 3, \\ \delta_1 &\sim IG(a_{z1}, b_{z1}), \\ \delta_2 &\sim IG(a_{z2}, b_{z2}). \end{aligned}$$

3.5 Additional assumptions

Define $\boldsymbol{\theta_1} = (\theta_{11}, \theta_{12}, \theta_{13})^T$, $\boldsymbol{\theta_2} = (\theta_{21}, \theta_{22}, \theta_{23})^T$, $\boldsymbol{\pi} = (\pi_{11}, \ldots, \pi_{I1}, \ldots, \pi_{14}, \ldots, \pi_{I4})$, $\mathbf{h} = (h_{11}, \ldots, h_{I1}, \ldots, h_{13}, \ldots, h_{I3})$, $\mathbf{u} = (u_{11}, \ldots, u_{I1}, \ldots, u_{14}, \ldots, u_{I4})$, and $\mathbf{v} = (v_{11}, \ldots, v_{1I}, \ldots, v_{13}, \ldots, v_{I3})$. The $\boldsymbol{\pi}$ and $\mathbf{u}$ are $4I \times 1$ vectors. The $\mathbf{h}$ and $\mathbf{v}$ are $3I \times 1$ vectors. In addition to the priors assumed for various parameters, the following prior independence assumptions are common in Bayesian modeling. Without the following independent assumptions, the prior distribution will be too complicated to work with. Note that the posterior distributions are not independent even we assume that their prior distributions are independent.

(i) Given $(\theta_{2j}, Z_{2i}, \delta_{2j})$, h_{ij} (or v_{ij}) is independent of $(\theta_{1j}, Z_{1i}, \delta_{1j}, \delta_1, \delta_2, \rho_1, \rho_2, \rho_3)$; where $i = 1, \cdots, I$ and $j = 1, 2, 3$.

(ii) Given $(\theta_{1j}, Z_{1i}, \delta_{1j})$, π_{ij} (or u_{ij}) is independent of $(\theta_{2j}, Z_{2i}, \delta_{2j}, \delta_1, \delta_2, \rho_1, \rho_2, \rho_3)$; where $i = 1, \cdots, I$ and $j = 1, 2, 3, 4$.

(iii) Given $(\rho_1, \rho_2, \rho_3, \delta_1, \delta_2)$, $\mathbf{Z}$ is independent of $(\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14}, \delta_{21}, \delta_{22}, \delta_{23},)$;

(iv) $(\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14}, \delta_{21}, \delta_{22}, \delta_{23}, \rho_1, \rho_2, \rho_3, \delta_1, \delta_2)$ are mutually independent.

Note that Assumption (3.5) basically assumes that the prior distribution of h_{ij} depends on only the parameters used in equation (5). Similarly, Assumption (3.5) assumes that the prior distribution of π_{ij} depends on only the parameters used in equation (4) or (8). Assumption (3.5) assumes that the prior distribution of $\mathbf{Z}$ depends on only the parameters used in equation (10).

4. BAYESIAN COMPUTATION

The unknown parameters in the model are $\boldsymbol{\pi}, \mathbf{h}$ (or equivalently $\mathbf{u}, \mathbf{v}$), $\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \mathbf{Z}_1, \mathbf{Z}_2, \rho_1, \rho_2, \rho_3, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14}, \delta_{21}, \delta_{22}, \delta_{23}, \delta_1$, and δ_2, a total number of $9I + 18$ parameters. In the case of MDHAS, $I = 114$, resulting in a total of 1044 parameters. As it is not feasible to obtain the closed form expression for a posterior distribution of such a large dimension, MCMC methods (Gilks et al., 1996) such as Gibbs sampling are used to evaluate the posterior distribution. The full conditional posterior distributions needed for the Gibbs sampling are given in Proposition 1 in the Appendix. The missing value y_{i41} is also simulated by the conditional distribution $(y_{i41} \mid \cdot)$ given in (ii) of Proposition 1.

A few special sampling techniques are used to improve the sampling efficiency. To sample δ_1 and δ_2, we use a transformation $\delta_k^* = \delta_k^{-1/2}$, $k = 1, 2$. The conditional densities of δ_1^*, δ_2^*, u_{ij}, and v_{ij} are log-concave. The proof is given in Proposition 2 in the Appendix. The adaptive rejection sampling method (Gilks and Wild, 1992) is used to sample from these log-concave conditional densities. The Metropolis-Hasting algorithm (Robert and Casella, 2004) is used to sample the conditional densities of ρ_1 and ρ_2 and the ratio of the two uniforms method (Wakefield et al., 1991) is used to sample the conditional density of ρ_3. Other conditional densities are standard distributions, such as normal, inverse gamma or multivariate normal. All of these can be easily generated. In the multivariate normal case, spectral decomposition (Johnson and Wichern, 1988) of the adjacency matrix $\mathbf{C}$ and Cholesky decomposition (Gelman et al., 2004) of variance matrix are used.

5. DATA ANALYSIS

5.1 Specification of hyperparameters

To implement the Gibbs sampling, the hyperparameters should be specified. Proper selection of hyperparameters is

one inherent difficulty with the parametric Bayes approach (Carlin and Louis, 2000). Recall that the means of logit of the empirical satisfaction rates for phase 1, 2, and 3 respondents are 1.4213, 1.0913, and 1.1818 respectively. Therefore we set $\mu_{11} = 1.4213$, $\mu_{12} = 1.0913$, and $\mu_{13} = 1.1818$.

Similarly, the means of logit of the empirical response rates for phase 1, 2, and 3 respondents are -0.8548, -1.4550, and -2.3488 respectively. Therefore we set $\mu_{21} = -0.8548$, $\mu_{22} = -1.4550$, and $\mu_{23} = -2.3488$.

For the values of τ_{1j} and τ_{2j}, $j = 1, 2, 3$, we assume $\tau_{11} = \tau_{12} = \tau_{13}$, and $\tau_{21} = \tau_{22} = \tau_{23}$. We have tried different values such as 4, 10 and 100 for τ_{1j} and for τ_{2j}, and the results are similar. The final results reported here are based on $\tau_{1j} = \tau_{2j} = 100$, $j = 1, 2, 3$. The estimates of θ_{1j} and θ_{2j} are very robust to the choices of (μ_{1j}, τ_{1j}), $j = 1, \ldots, 4$ and (μ_{2j}, τ_{2j}), $j = 1, 2, 3$. So are the estimated satisfaction rates and response rates.

It is difficult to set the hyperparameters of the variance components. The choice of (a_{1j}, b_{1j}), $j = 1, \ldots, 4$ and (a_{2j}, b_{2j}), $j = 1, 2, 3$ can have some influence on the estimates of δ_{1j}, $j = 1, \ldots, 4$ and δ_{2j}, $j = 1, 2, 3$, but not the estimated satisfaction rates and response rates. Therefore we begin with a noninformative constant prior for the parameters δ_{1j}, $j = 1, \ldots, 4$, δ_{2j}, $j = 1, 2, 3$, δ_1, and δ_2. We then use the posterior distributions obtained based on the constant prior to determine the hyperparameters. Let m_0 and v_0 be the posterior mean and variance of the target parameters respectively. We choose hyperparameters for $IG(a, b)$ such that

$$\frac{b}{a-1} = 2m_0, \qquad \frac{b^2}{(a-1)^2(a-2)} = 200v_0.$$

Solving a and b for the above equations, the resulting hyperparameters values are listed in equations (12) to (15) except (a_{14}, b_{14}). Based on the assumption that the satisfaction rate given nonresponse is similar to the satisfaction rate given response to the third phase, the same hyperparameters are used for δ_{14} as for δ_{13}, i.e., set $(a_{14}, b_{14}) = (a_{13}, b_{13})$.

To compare the estimates under the proposed priors with the ones under the flat priors, see the scatter plots in Figures 1 and 2. Figure 1 displays the comparisons on the estimates of conditional satisfaction rates. The estimates under proposed priors are plotted along x-axis, and the estimates under the flat priors are plotted along y-axis. The left column is for comparing the estimates, and the right column is for comparing the standard deviations of those estimates. The first three rows are for mailing phase 1 to 3 respectively, and the fourth row is for the non-respondents. Note that the estimates and their standard deviations are almost the same for mailing phase 1 to 3. The estimates under proposed priors are slightly larger than the ones under the flat priors for the non-respondents, while the corresponding estimated standard deviations are much smaller. This is because that the proposed priors borrow information from
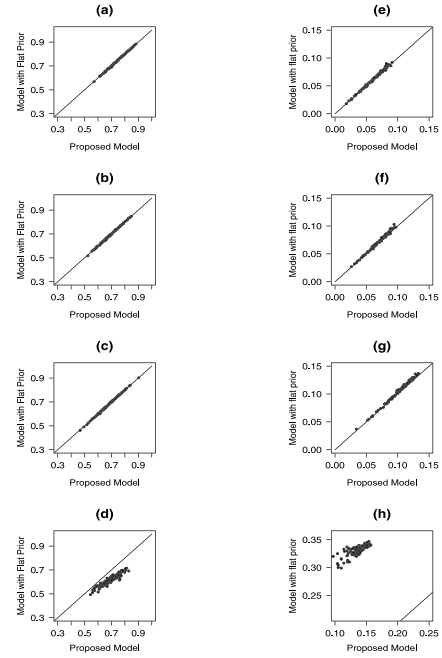

Figure 1. Scatterplots of conditional satisfaction rate estimates from proposed priors on x-axis vs ones with flat priors on y-axis, (a)–(d): $\hat{\pi}_{ij}$, $j = 1, \ldots, 4$; (e)–(h): standard deviations of the estimates for $j = 1, \ldots, 4$.
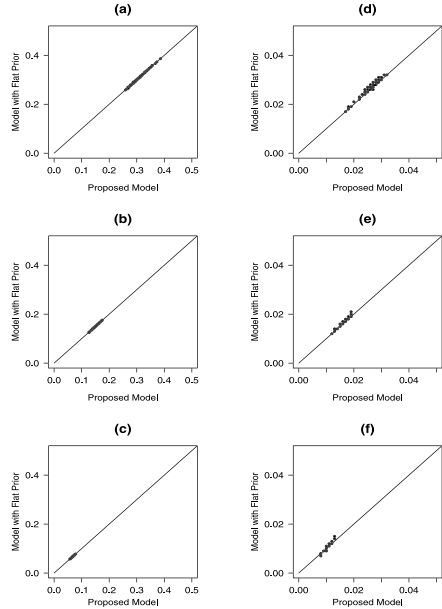

Figure 2. Scatterplots of response rate estimates from proposed priors on x-axis vs ones with flat priors on y-axis, (a)–(c): $\hat{q}_{ij}$, $j = 1, 2, 3$; (d)–(f): the corresponding standard deviations for $j = 1, 2, 3$.

mailing phase 3. There is no data to estimate δ_{14} if there is no information shared between mailing phase 3 and non-respondents. The similar comparisons on the estimates of

response rates are showed in Figure 2. Again, the estimates and their standard deviations are almost the same for mailing phase 1 to 3.

In summary, the results reported in this paper are based on the following specifications.

- The hyperparameters for θ_{1j}, $j = 1, 2, 3$ as

$$
\begin{aligned}
(\mu_{11}, \tau_{11}) &= (1.4213, 100), \\
(\mu_{12}, \tau_{12}) &= (1.0913, 100), \\
(\mu_{13}, \tau_{13}) &= (1.1818, 100);
\end{aligned}
$$

- The hyperparameters for θ_{2j}, $j = 1, 2, 3$ as

$$
\begin{aligned}
(\mu_{21}, \tau_{21}) &= (-0.8548, 100), \\
(\mu_{22}, \tau_{22}) &= (-1.4550, 100), \\
(\mu_{23}, \tau_{23}) &= (-2.3488, 100);
\end{aligned}
$$

- The hyperparameters for δ_{1j}, $j = 1, \ldots, 4$ as

$$
\begin{aligned}
(a_{11}, b_{11}) &= (2.0303, 0.1267), \\
(a_{12}, b_{12}) &= (2.0066, 0.1303), \\
(a_{13}, b_{13}) &= (2.0096, 0.4925), \\
(12) \quad (a_{14}, b_{14}) &= (2.0096, 0.4925);
\end{aligned}
$$

- The hyperparameters for δ_{2j}, $j = 1, 2, 3$ as

$$
\begin{aligned}
(a_{21}, b_{21}) &= (2.0247, 0.0143), \\
(a_{22}, b_{22}) &= (2.0339, 0.0428), \\
(13) \quad (a_{23}, b_{23}) &= (2.0298, 0.0741);
\end{aligned}
$$

- The hyperparameters for δ_1 as

$$
(14) \quad (a_{z1}, b_{z1}) = (2.2256, 0.3260);
$$

- The hyperparameters for δ_2 as

$$
(15) \quad (a_{z2}, b_{z2}) = (2.1391, 0.0415).
$$

5.2 Bayesian estimation

The Gibbs sampling trace plots are examined for all model parameters and some randomly selected spatial effects Z_{1i} and Z_{2i} and they converge after 1,000 iterations. Final results presented here are based on 30,000 Gibbs cycles after burn-in of the initial 2,000 iterations. The posterior distributions are approximated by MCMC samples in Figure 3, 4, and 5. The 90% credible intervals are given in Table 4 along with the posterior means, medians, and standard deviations.

The Response Phase Effects. The posterior distributions of θ_{11}, θ_{12}, and θ_{13} are given in Figure 3 (a), (b), and (c), respectively. The 90% credible intervals of $\theta_{11}, \theta_{12}, \theta_{13}$ are all positive. Posterior means are close to posterior medians. The center of 90% credible intervals are very close to the posterior means as well. The concentrations of the posterior distributions are all away from zero to the right. To
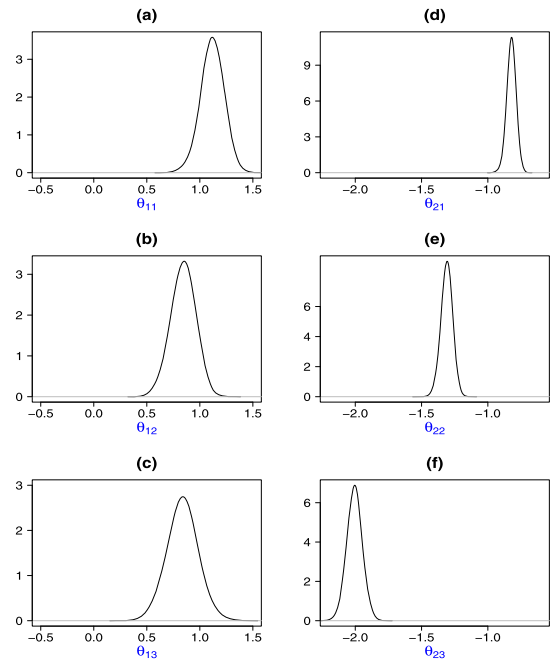

Figure 3. Marginal posterior densities of $(\theta_{11}, \theta_{12}, \theta_{13}, \theta_{21}, \theta_{22}, \theta_{23})$.
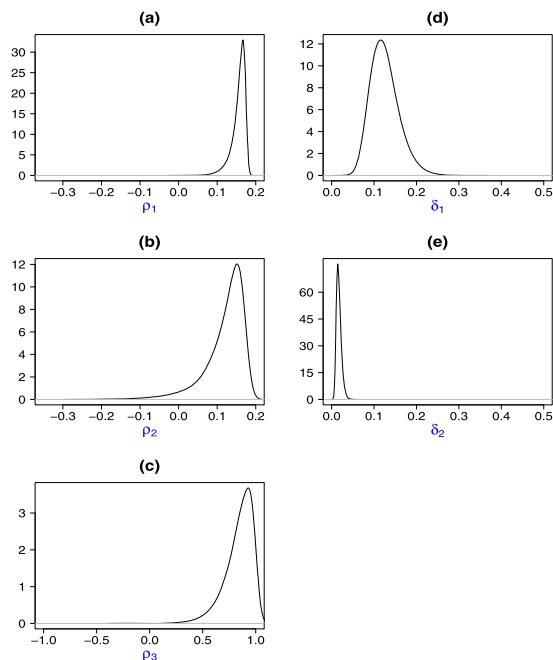

Figure 4. Marginal posterior densities of (ρ_1, ρ_2, ρ_3) and (δ_1, δ_2).

compare the response phase effects θ_{1j} on conditional satisfaction rates π_{1j}, we see that the posterior distribution of θ_{11} shifts to the right comparing with the distributions of θ_{12} and θ_{13}. The posterior distributions of θ_{12} and θ_{13} are similar, but 90% credible interval of θ_{13} is slightly wider.

Table 4. 90% credible intervals, posterior means, medians, and standard deviations of model parameters

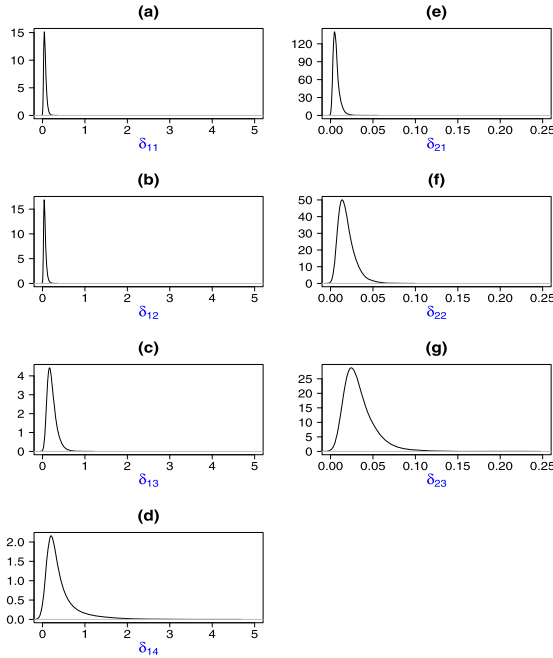| Parameter	90% credible interval	posterior mean (median)	STD
$\hat{\theta}_{11}$	(0.9375, 1.2904)	1.1178 (1.1191)	0.1075
$\hat{\theta}_{12}$	(0.6505, 1.0252)	0.8430 (0.8462)	0.1139
$\hat{\theta}_{13}$	(0.6008, 1.0395)	0.8373 (0.8378)	0.1431
$\hat{\theta}_{21}$	(-0.8790, -0.7673)	-0.8222 (-0.8216)	0.0343
$\hat{\theta}_{22}$	(-1.3819, -1.2413)	-1.3098 (-1.3089)	0.0426
$\hat{\theta}_{23}$	(-2.0986, -1.9117)	-2.0061 (-2.0059)	0.0567
$\hat{\rho}_1$	(0.1203, 0.1736)	0.1590 (0.1613)	0.0179
$\hat{\rho}_2$	(0.0183, 0.1700)	0.1205 (0.1355)	0.0509
$\hat{\rho}_3$	(0.5920, 0.9882)	0.8431 (0.8722)	0.1280
$\hat{\delta}_1$	(0.0794, 0.1843)	0.1251 (0.1215)	0.0325
$\hat{\delta}_2$	(0.0098, 0.0279)	0.0171 (0.0162)	0.0057
$\hat{\delta}_{11}$	(0.0232, 0.1212)	0.0600 (0.0529)	0.0315
$\hat{\delta}_{12}$	(0.0233, 0.1181)	0.0575 (0.0498)	0.0310
$\hat{\delta}_{13}$	(0.0910, 0.4253)	0.2193 (0.1964)	0.1071
$\hat{\delta}_{14}$	(0.1000, 1.2157)	0.4155 (0.2855)	0.3928
$\hat{\delta}_{21}$	(0.0026, 0.0137)	0.0067 (0.0058)	0.0037
$\hat{\delta}_{22}$	(0.0076, 0.0374)	0.0188 (0.0167)	0.0096
$\hat{\delta}_{23}$	(0.0132, 0.0671)	0.0336 (0.0295)	0.0178

Figure 5. Marginal posterior densities of (δ_{11}, δ_{12}, δ_{13}, δ_{14}, δ_{21}, δ_{22}, δ_{23}).

which indicates that response rates of the three phases are very different. The response rates to the first phase are the highest while the response rates to the third phase are the lowest.

Spatial Dependencies. The posterior distributions of ρ_1, ρ_2, and ρ_3 are shown in Figure 4 (a), (b), and (c), respectively. The posterior mean of ρ_1 is 0.1590 with a standard deviation of 0.0179. The posterior mean of ρ_2 is 0.1205 with a standard deviation of 0.0509. Compared to the possible range of $(-0.3457, 0.1756)$, the 90% credible interval $(0.1203, 0.1736)$ and the concentration of the posterior distribution of ρ_1 indicates a strong spatial dependency for the conditional satisfaction rates (Figure 4 (a)), and the 90% credible interval $(0.0183, 0.1700)$ and the concentration of the posterior distribution of ρ_2 indicate a moderate spatial dependency for phase specific response rates (Figure 4 (b)). The posterior mean of ρ_3 is 0.8431 with a standard deviation of 0.1280. Compared to its possible range of $(-1, 1)$, the 90% credible interval and the concentration of the posterior distribution of ρ_3 indicate that spatial dependency between the conditional satisfaction rates and the conditional response rates is also strong. Thus the "bridging" parameter ρ_3 in $(\mathbf{Z}_1, \mathbf{Z}_2)$ shows a certain degree of additional information borrowing in the estimates of conditional satisfaction rates and phase specific response rates. This indicates that counties with higher conditional satisfaction rates tend to have higher response rates to the 2001 MDHAS.

Variance Components. The posterior distributions of δ_1 and δ_2 are shown in Figure 4 (d) and (e) respectively. Both posterior distributions are approximately symmetric about their means. The posterior mean of δ_1 is 0.1251 with a standard deviation of 0.0325 while the posterior mean of δ_2 is
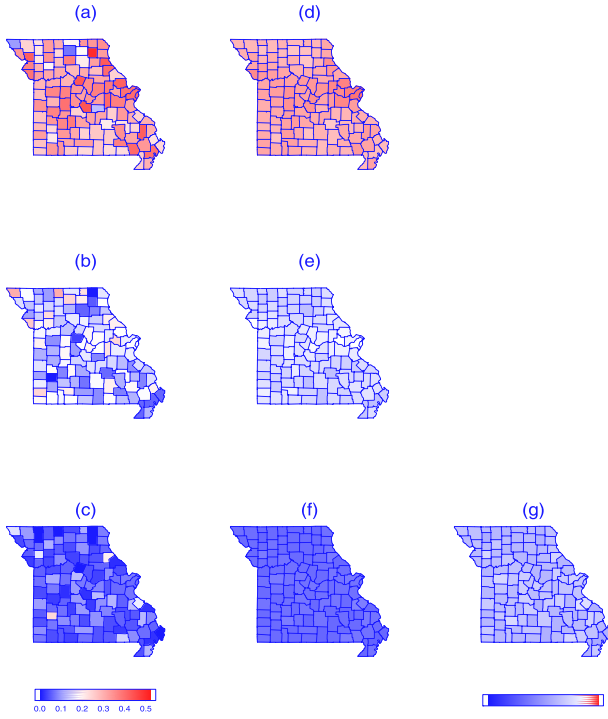
The posterior distributions of θ_{21}, θ_{22}, and θ_{23} are given in Figure 3 (d), (e), and (f), respectively. The 90% credible intervals are all negative. Posterior means are close to the posterior medians. The center of 90% credible intervals are very close to the posterior means as well. The concentrations of the posterior distributions are all away from zero to the left. There is little overlap in the posterior distributions,

Figure 6. Estimates of response rates. (a)–(c): Frequency estimates $\hat{q}^F_{ij}$, $(j = 1, 2, 3)$; (d)–(f): Bayesian estimates $\hat{q}_{ij}$, $(j = 1, 2, 3)$. (g): $\hat{q}_{i3}$ on a different scale.

0.0171 with a standard deviation of 0.0057. The 90% credible intervals are $(0.0794, 0.1843)$ and $(0.0098, 0.0279)$ respectively. All the above indicate that the spatial variation among conditional satisfaction rates is larger than the spatial variation among phase specific response rates.

The posterior distributions of δ_{1j} and δ_{2j} are shown in Figure 5 (a)–(g), respectively. All the distributions are right skewed. The posterior means, medians, 90% credible intervals, and standard deviations are given in Table 4. In general δ_{1j} $(j = 1, \ldots, 4)$ are larger than δ_{2j} $(j = 1, 2, 3)$. The posterior mean or median of δ_{11} and δ_{12} are about the same. The posterior mean or median of δ_{13} is larger than the posterior mean or median of δ_{11} and δ_{12}. The posterior mean or median of δ_{14} is the largest among δ_{1j}, $j = 1 \ldots, 4$. The posterior mean or median of δ_{21} is the smallest while the posterior mean of median of δ_{23} is the largest among δ_{2j}, $j = 1, 2, 3$. All these results indicate that the variances of the phase-county effects are larger for the conditional satisfaction rates than that of the phase-county effects for the phase specific response rates.

Response Rates by County of Residence. The frequency estimates of the survey response rates, $\hat{q}^F_{ij} = n_{ij}/m_{i1}$ for $j = 1, 2, 3$ are shown in Figure 6 (a)–(c). The model-based Bayesian estimates of conditional response rates are $\hat{h}_{ij}$ for $j = 1, 2, 3$. Using the relationship of (2), the model-based Bayesian estimates of the survey response rates, $\hat{q}_{ij}$ for $j = 1, 2, 3$ can be obtained easily. They are plotted in

Figure 6 (d)–(f). Clearly, the response rates are the highest for the first phase and the lowest for the third phase. This is true for both the frequency and model estimates. However, the Bayesian model gives more stable (smoother) estimates than the frequency method, especially for counties with small sample sizes. The Bayesian estimates borrow information from surrounding counties and the individual estimates are pulled toward their local averages. The frequency estimates are unstable with small sample sizes. For example, for an increase of 1 respondent in a county with a sample size of 8, the response rate will increase 12.5%.

Note that the Bayesian estimates $\hat{q}_{i3}$ seem to be over smoothed comparing with $\hat{q}^F_{i3}$, see Figure 6 (c) and (f). We have checked this and found that the over smoothing is not related to the choice of priors. Figure 2 (c) shows that the estimates of response rates are almost the same under the flat priors and proposed priors, which are two very different priors. We also fit the model with a different prior on ρ_2. It does not change the results on q_{i3}. The over smoothing is actually related to 1) much smaller counts (respondents) in phase 3 than in phase 1 and 2; 2) the spatial structure on response rates (all 3 phases are assumed the same spatial structure). So the spatial pattern of estimated $\hat{q}_{i3}$ is dominated by the spatial pattern of the estimated $\hat{q}_{i1}$ and $\hat{q}_{i2}$, see Figure 6 (g). For this application, we do not have evidence to assume the spatial structure of q_{i3} is different from the spatial structure of q_{i1} or q_{i2}. The model verification (see session 5.3) also indicates this is a reasonable model. In other cases, one may want to assume different spatial structures for different phases. But, how to share information among different phases and between satisfaction rate and response rate is a question needed for further study.

Satisfaction Rates by County of Residence. The frequency estimates of the conditional satisfaction rates given the response phase, $\hat{\pi}^F_{ij} = y_{ij}/n_{ij}$ for $j = 1, 2, 3$, are shown in Figure 7 (a)–(c). Note that the frequency estimates of conditional satisfaction rates given nonresponse are not available. The Bayesian estimates of the satisfaction rates given response status, $\hat{\pi}_{ij}$ for $j = 1, \ldots, 4$, are shown in Figure 7 (d)–(g). At the county level, frequency estimates range from 0.3333 to 1.0 for the first response phase, and 0 to 1.0 for the second and third response phases. In the third response phase, there are 28 counties with conditional satisfaction rates as 100%. We don't expect this to be true. In fact all these counties have small sample sizes. For example, Reynolds county had 2 respondents in the third phase, and both respondents were satisfied with MDC's deer management. On the other hand, the model-based Bayesian estimates of conditional satisfaction rates are more reasonable, as they range from 0.5710 to 0.8830 for the first response phase, and from 0.5240 to 0.8480 for the second, and from 0.4670 to 0.9030 for the third response phase. The conditional satisfaction rate estimates for nonrespondents range from 0.5440 to 0.8290. Several counties have no data for given response phases as shown by counties shaded in the map ((Figure 7 (b) and (c)). In these cases, there are no
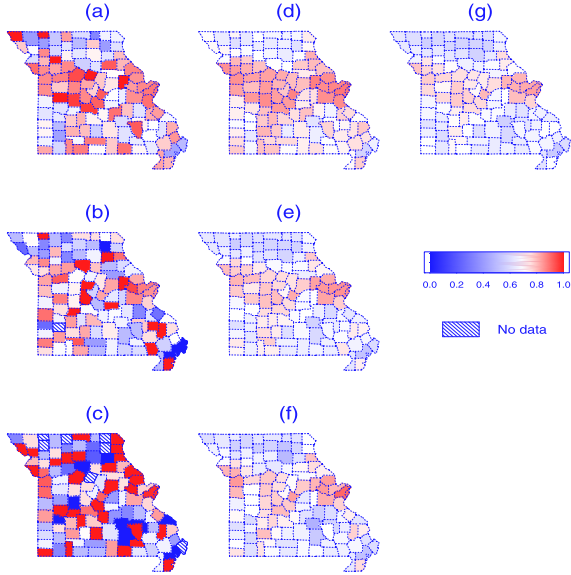
Figure 7. Maps of naive frequency estimates of $\hat{\pi}_{ij}^F = y_{ij}/n_{ij}$ and Bayesian estimates of $\hat{\pi}_{ij}$. (a)–(c): Frequency estimates of $\hat{\pi}_{ij}^F$, $(j = 1, 2, 3)$ for response phase 1, 2, and 3, respectively; (d)–(g): Bayesian estimates of $\hat{\pi}_{ij}$, $(j = 1, \ldots, 4)$ for response phase 1, 2, 3, and survey nonrespondents, respectively.

frequency estimates, but Bayesian estimates are available. Under the Bayesian model, individual counties borrow information from surrounding areas and estimates are pulled toward their local averages (Figure 7 (d)–(g)).

Again our interest is to estimate the marginal satisfaction rates p_i at the county level. From Table 3, marginal satisfaction rates for county i, $i = 1, \ldots, I$ are

$$(16) \qquad \hat{p}_i = \hat{\pi}_{i1}\hat{q}_{i1} + \hat{\pi}_{i2}\hat{q}_{i2} + \hat{\pi}_{i3}\hat{q}_{i3} + \hat{\pi}_{i4}\left(1 - \sum_{j=1}^{3}\hat{q}_{ij}\right).$$

The frequency marginal satisfaction rates are not available because there are no data for nonrespondents. To make a fair comparison, for each county we compare the following two estimates as well.

- Conditional satisfaction rates given response to the survey

$$\hat{\pi}_i = \frac{\hat{\pi}_{i1}\hat{q}_{i1} + \hat{\pi}_{i2}\hat{q}_{i2} + \hat{\pi}_{i3}\hat{q}_{i3}}{\hat{q}_{i1} + \hat{q}_{i2} + \hat{q}_{i3}}, \quad i = 1, \ldots, I.$$

- The frequency counterparts

$$\hat{\pi}_i^F = \frac{y_{i1} + y_{i2} + y_{i3}}{n_{i1} + n_{i2} + n_{i3}}, \quad i = 1, \ldots, I.$$

The $\hat{\pi}_i^F$, $\hat{\pi}_i$, and $\hat{p}_i$ are shown in maps in Figure 8 (a)–(c). The overall spatial patterns are similar. The frequency estimates of conditional satisfaction rates given response range from 0.36 to 1.0 while the model-based estimates range from
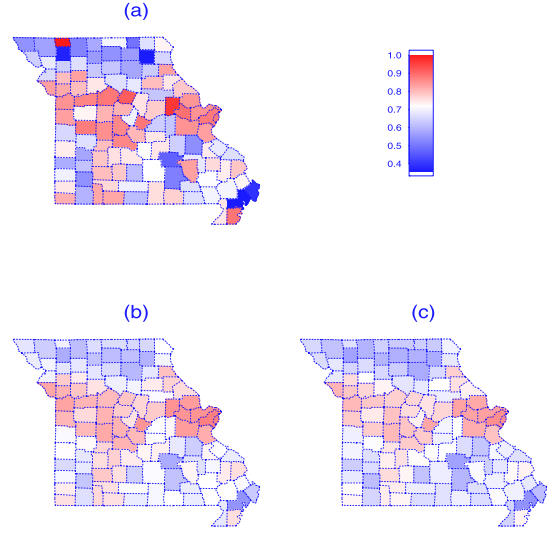
Figure 8. All phase combined estimates. (a): $\hat{\pi}_i^F$, the naive frequency estimate of all phase combined conditional satisfaction rates; (b): $\hat{\pi}_i$, the Bayesian estimates of all phase combined conditional satisfaction rates; (c): $\hat{p}_i$, Bayesian estimates of marginal satisfaction rates.
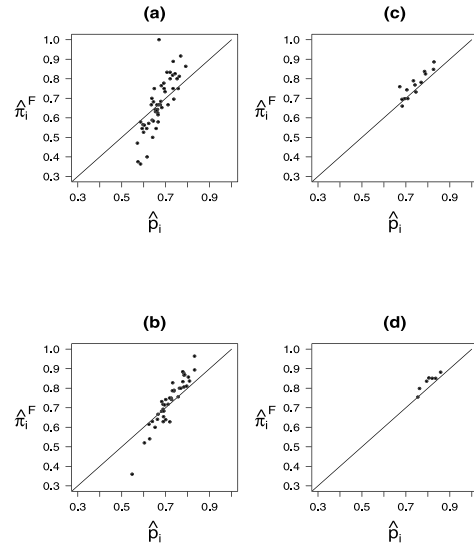

Figure 9. Comparison of estimated satisfaction rates from naive frequency method, $\hat{\pi}_i^F$ and Bayesian method, $\hat{p}_i$ when sample sizes (a): $n_i < 25$; (b): $25 \le n_i < 50$; (c): $50 \le n_i < 100$; (d): $n_i \ge 100$.

0.5475 to 0.8755. Again we can see that Bayesian estimates are pulled toward their local averages and are smoother than frequency estimates. In Figure 9 (a)–(d) we present 4 scatter plots of $\hat{\pi}_i^F$ vs $\hat{p}_i$ by 4 categories of sample sizes or number of respondents n_i, where $n_i = n_{i1} + n_{i2} + n_{i3}$. It is clear that the difference between Bayesian and frequency estimates is large when sample sizes are small.

Phase j	$\hat{q}_j^F$	$\hat{q}_j$	$\hat{\pi}_j^F$	$\hat{\pi}_j$
1	0.3196	0.3196	0.7885	0.7884
2	0.1524	0.1525	0.7418	0.7417
3	0.0662	0.0663	0.7312	0.7309
nonresponse	N.A	N.A	N.A	0.7097

5.3 Model verification and statewide estimates

When the sample size is large, the frequency estimate is a good estimate. Therefore, we can verify the model-based estimates by comparing them with the frequency estimates at state level. Let $\hat{q}_j$ and $\hat{q}_j^F$ be the statewide Bayesian and frequency estimates of response rates for response phase j. Let $\hat{\pi}_j$ and $\hat{\pi}_j^F$ be the statewide Bayesian and frequency estimates of conditional satisfaction rates given response for response phase j. Recall that for each county i, m_{i1} are the total number of hunters who received the survey in the first phase. n_{ij} for $j = 1, 2, 3$, are the total number of respondents for the first, second, and third phase, respectively. n_{i4} is the number of nonrespondents (Table 1). Let $m_{i2} = m_{i1} - n_{i1}$ and $m_{i3} = m_{i2} - n_{i2}$, then m_{i2} and m_{i3} are the total number of hunters who received surveys for the second and third phases, respectively. Thus,

$$\hat{q}_j = \frac{\sum_{i=1}^I m_{ij}\hat{q}_{ij}}{\sum_{i=1}^I m_{ij}}, \quad j = 1, 2, 3,$$

$$\hat{q}_j^F = \frac{\sum_{i=1}^I n_{ij}}{\sum_{i=1}^I m_{i1}}, \quad j = 1, 2, 3,$$

$$\hat{\pi}_j = \frac{\sum_{i=1}^I n_{ij}\hat{\pi}_{ij}}{\sum_{i=1}^I n_{ij}}, \quad j = 1, \ldots, 4,$$

$$\hat{\pi}_j^F = \frac{\sum_{i=1}^I y_{ij}}{\sum_{i=1}^I n_{ij}}, \quad j = 1, 2, 3.$$

The results are given in Table 5.

Combined statewide Bayesian estimates of the response rate for all phases, $\hat{q}$, and the conditional satisfaction rate given response, $\hat{\pi}$, are defined as

$$\hat{q} = \hat{q}_1 + \hat{q}_2 + \hat{q}_3,$$

$$\hat{\pi} = \frac{\sum_{j=1}^3 \hat{\pi}_j\hat{q}_j}{\hat{q}}.$$

Combined statewide frequency estimates of the response rate for all phases, $\hat{q}^F$, and the conditional satisfaction rate given response, $\hat{\pi}^F$, are defined as

$$\hat{q}^F = \frac{\sum_{i=1}^I \sum_{j=1}^3 n_{ij}}{\sum_{i=1}^I m_{i1}}.$$

$$\hat{\pi}^F = \frac{\sum_{i=1}^I \sum_{j=1}^3 y_{ij}}{\sum_{i=1}^I \sum_{j=1}^3 n_{ij}}.$$

The marginal statewide satisfaction rate is defined as

$$\hat{p} = \frac{\sum_{i=1}^I m_{i1}\hat{p}_i}{\sum_{i=1}^I m_{i1}},$$

where $\hat{p}_i$ is defined in (16).

The combined statewide estimates of response rates for all phases are $\hat{q} = 0.5384$ and $\hat{q}^F = 0.5382$ from Bayesian and frequency methods, respectively. The all phase combined statewide estimates of conditional satisfaction rates given response are $\hat{\pi} = 0.768092$ and $\hat{\pi}^F = 0.768226$. It is a surprise to see that the Bayesian estimates match the frequency estimates almost exactly. This indicates that the Bayesian model fits the given data reasonably.

The marginal statewide Bayesian estimate of satisfaction rate is $\hat{p} = 0.741435$ which is about 2.7% less than the frequency estimate of conditional satisfaction rate given response. In other words, without adjusting for nonresponse, we would overestimate the number of satisfied hunters by more than 9,000 since there were 357,552 Missouri hunters who bought firearm deer hunting permits in 2001.

6. DISCUSSION

In this article, we have developed a Bayesian hierarchical model to estimate the satisfaction rates at subdomains without covariates via modeling the phase specific response rates and conditional satisfaction rates given response statuses and accounting for small sample size and survey nonresponse. In this pattern mixture approach, there are no data to estimate the conditional satisfaction rates given nonresponse. We define a direct relationship to relate conditional satisfaction rates given nonresponse to that of the third phase response through a linear mixed model. Phase specific response rates and satisfaction rates given response statuses are then calculated and combined to get the marginal satisfaction rates p_i.

We find that there are spatial dependencies in conditional satisfaction rates and phase specific response rates among hunters' Missouri county of residence. The response phases affect the conditional satisfaction and phase specific response rates. The effects on response rates are more prominent than on the conditional satisfaction rates. In general, the first phase has the highest conditional satisfaction rates and response rates while the third phase has the lowest conditional satisfaction rates as well as response rates. But the response rates for the three phases are different while the conditional satisfaction rates for the second and third phase are similar.

At the county level, model-based estimates were more reasonable than their frequency counterparts. For example, frequency estimates range from 0.36 to 1.0. We do not expect any counties to have a hunter satisfaction rate of 100%. Under the Bayesian model, the county borrowed informa-

tion from surrounding areas and estimates are pulled toward their local average, especially for counties with no data or very small sample sizes. The model also gives an estimate of conditional satisfaction rates given nonresponse for each county.

The model-based statewide estimates of response rates and conditional satisfaction rates given response to phases 1, 2 and 3 are very close to their frequency counterparts. This is one way to evaluate the model-based estimates because frequency estimates should be reasonable when sample sizes are large. Overall we found that at the statewide level the conditional satisfaction rates given nonresponse are lower than the conditional satisfaction rates given response. The marginal satisfaction rates are lower than the conditional satisfaction rates given response. If we ignored nonresponse, we would overestimate the number of satisfied hunters by more than 9,000 statewide.

Gao et al. (2014) analyzed this 2001 MDHAS using a bivariate Bayesian hierarchical linear mixed model to estimate hunters' satisfaction rate and response rate. Both estimated rates were then aggregated at hunters' county of residence and at the state level respectively. In addition to random spatial effects to account for dependency among subdomains, they also included auxiliary covariates of hunters, such as age and number of deer harvested. The overall estimated statewide satisfaction rate was 0.7366 comparing to an estimate of 0.7414 given in this article. The phase specific estimates of satisfaction rates were 0.7750, 0.7365, and 0.7100 for the first mailing phase, late mailing phase (the second and third mailing combined), and nonresponse respectively. Comparing to the results presented in Table 5, the estimates from the two different approaches are very close. Auxiliary information is not always available or hard to obtain for surveys. In this article, we provide a way to utilize the mailing phases and spatial data to assess and adjust survey nonresponse when the main objective is to estimate subdomain and overall satisfaction and response rates. The model presented here is easier to implement than the one provided in Gao et al. (2014) which used hunter level data and had 21,947 parameters in the model. Here we use subdomain level data. The model in section 3 has 1,044 parameters and takes less time to run.

The general approach presented here is applicable to other surveys. However, one should be careful when specifying link function in the linear mixed model. One also should check assumptions for each survey carefully when linking nonresponse to response through multiple survey phases. One can also check if there are other auxiliary information about survey respondents and nonrespondents. That auxiliary information may also be incorporated into the models. Finally, it is best to collect satisfaction data on nonresponse directly to further validate the model.

ACKNOWLEDGMENT

APPENDIX

Proposition 1. *Let* $\mathbf{y}^* = (y_{141}, \ldots, y_{I41})$, $\mathbf{y}^o = (y_{111}, y_{121}, y_{131}, \ldots, y_{I11}, y_{I21}, y_{I31})$, *and* $\mathbf{n} = (n_{11}, n_{12}, n_{13}, n_{14}, \ldots, n_{I1}, n_{I2}, n_{I3}, n_{I4},)$. *Here* $\mathbf{y}^*$ *is a* $I \times 1$ *vector,* $\mathbf{y}^o$ *is a* $3I \times 1$ *vector, and* $\mathbf{n}$ *is a* $4I \times 1$ *vector. Let* $[\Omega \mid others; data]$ *denote the conditional density of* Ω *given all other parameters and* $(\Omega \mid others; data)$ *denote the conditional distribution of* Ω *given all other parameters, where data are* $(\mathbf{y}^o, \mathbf{n})$. *The full conditional posterior distributions are as follows.*

(i) $[u_{ij} \mid others; data] = [u_{ij} \mid \boldsymbol{\theta_1}, \mathbf{Z}_1, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14}, \mathbf{y}^*;$ $\mathbf{y}^o, \mathbf{n}]$ *for* $i = 1, \ldots, I$, $j = 1, 2, 3$ *is proportional to*

$$\exp\left\{ u_{ij}y_{ij1} - n_{ij}\log(1+e^{u_{ij}}) - \frac{(u_{ij}-\theta_{1j}-z_{1i})^2}{2\delta_{1j}} \right\}.$$

$[u_{i4} \mid others; data] = [u_{i4} \mid \boldsymbol{\theta_1}, \mathbf{Z}_1, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14}, \mathbf{y}^*;$ $\mathbf{y}^o, \mathbf{n}]$ *for* $i = 1, \ldots, I$ *is proportional to*

$$\exp\left\{ u_{i4}y_{i41} - n_{i4}\log(1+e^{u_{i4}}) - \frac{(u_{i4}-\theta_{13}-z_{1i})^2}{2\delta_{14}} \right\}.$$

(ii) $(y_{i41} \mid others; data) \sim Binomial(n_{i4}, \pi_{i4})$, *where* $\pi_{i4} = \dfrac{e^{u_{i4}}}{1+e^{u_{i4}}}$.

(iii) $[v_{ij} \mid others; data] = [v_{ij} \mid \boldsymbol{\theta_2}, \mathbf{Z}_2, \delta_{21}, \delta_{22}, \delta_{23}; \mathbf{n}]$ *for* $i = 1, \ldots, I$, $j = 1, 2, 3$ *is proportional to*

$$\exp\left\{ v_{ij}n_{ij} - m_{ij}\log(1+e^{v_{ij}}) - \frac{(v_{ij}-\theta_{2j}-z_{2i})^2}{2\delta_{2j}} \right\},$$

where $m_{i1} = \sum_{j=1}^{4} n_{ij}$, $m_{i2} = m_{i1} - n_{i1}$, $m_{i3} = m_{i2} - n_{i2}$.

(iv) $(\theta_{1j} \mid others; data) = (\theta_{1j} \mid \mathbf{u}, \mathbf{Z}_1, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14})$ *for* $j = 1, 2$ *is*

$$N\left(\frac{\tau_{1j}\sum_{i=1}^{I}(u_{ij}-z_{1i}) + \delta_{1j}\mu_{1j}}{I\tau_{1j}+\delta_{1j}}, \frac{\tau_{1j}\delta_{1j}}{I\tau_{1j}+\delta_{1j}} \right).$$

$(\theta_{13} \mid others; data) = (\theta_{13} \mid \mathbf{u}, \mathbf{Z}_1, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14})$ *is* $N(\mu, \sigma^2)$, *where*

$$\mu = \frac{\delta_{14}\tau_{13}\sum_{i=1}^{I}(u_{i3}-z_{1i}) + \delta_{13}\tau_{13}\sum_{i=1}^{I}(u_{i4}-z_{1i})}{I\tau_{13}(\delta_{13}+\delta_{14})+\delta_{13}\delta_{14}}$$
$$+ \frac{\mu_{13}\delta_{13}\delta_{14}}{I\tau_{13}(\delta_{13}+\delta_{14})+\delta_{13}\delta_{14}},$$
$$\sigma^2 = \frac{\delta_{13}\delta_{14}\tau_{13}}{I\tau_{13}(\delta_{13}+\delta_{14})+\delta_{13}\delta_{14}}.$$

(v) $(\theta_{2j} \mid others; data) = (\theta_{2j} \mid \mathbf{v}, \mathbf{Z}_2, \delta_{21}, \delta_{22}, \delta_{23})$ for $j = 1, 2, 3$ is

$$N\left(\frac{\tau_{2j}\sum_{i=1}^{I}(v_{ij} - z_{2i}) + \delta_{2j}\mu_{2j}}{I\tau_{2j} + \delta_{2j}}, \frac{\tau_{2j}\delta_{2j}}{I\tau_{2j} + \delta_{2j}}\right).$$

(vi) $(\mathbf{Z}_1 \mid others; data) = (\mathbf{Z}_1 \mid \mathbf{u}, \boldsymbol{\theta_1}, \mathbf{Z}_2, \delta_{11}, \delta_{12}, \delta_{13}, \delta_{14}, \rho_1, \rho_2, \rho_3, \delta_1, \delta_2)$ is

$$N\left(\mathbf{G}_1^{-1}\left[\frac{\rho_3}{(1 - \rho_3^2)}\frac{\mathbf{B}_1^{\frac{1}{2}}\mathbf{B}_2^{\frac{1}{2}}\mathbf{Z}_2}{\sqrt{\delta_1\delta_2}} + \mathbf{C}_1\right], \mathbf{G}_1^{-1}\right),$$

where $\mathbf{C}_1 = (c_{11}, \ldots, c_{1I})^T$, $c_{1i} = \sum_{j=1}^{3}\frac{u_{ij} - \theta_{1j}}{\delta_{1j}} + \frac{u_{i4} - \theta_{13}}{\delta_{14}}$, $i = 1, \ldots, I$, and

$$\mathbf{G}_1 = \sum_{j=1}^{4}\frac{1}{\delta_{1j}}\boldsymbol{I}_I + \frac{1}{(1 - \rho_3^2)\delta_1}\mathbf{B}_1.$$

(vii) $(\mathbf{Z}_2 \mid others; data) = (\mathbf{Z}_2 \mid \mathbf{v}, \boldsymbol{\theta_2}, \mathbf{Z}_1, \delta_{21}, \delta_{22}, \delta_{23}, \rho_1, \rho_2, \rho_3, \delta_1, \delta_2)$

$$N\left(\mathbf{G}_2^{-1}\left[\frac{\rho_3}{(1 - \rho_3^2)}\frac{\boldsymbol{B}_2^{\frac{1}{2}}\boldsymbol{B}_1^{\frac{1}{2}}\mathbf{Z}_1}{\sqrt{\delta_1\delta_2}} + \mathbf{C}_2\right], \mathbf{G}_2^{-1}\right),$$

where $\mathbf{C}_2 = (c_{21}, \ldots, c_{2I})^T$, $c_{2i} = \sum_{j=1}^{3}\frac{v_{ij} - \theta_{2j}}{\delta_{2j}}$, $i = 1, \ldots, I$, and

$$\mathbf{G}_2 = \sum_{j=1}^{3}\frac{1}{\delta_{2j}}\boldsymbol{I}_I + \frac{1}{(1 - \rho_3^2)\delta_2}\mathbf{B}_2.$$

(viii) $(\delta_{1j} \mid others; data) = (\delta_{1j} \mid \mathbf{u}, \boldsymbol{\theta_1}, \mathbf{Z}_1)$ for $j = 1, \ldots, 3$ is

$$IG\left(a_{1j} + \frac{I}{2}, b_{1j} + \frac{1}{2}\sum_{i=1}^{I}(u_{ij} - \theta_{1j} - z_{1i})^2\right).$$

$(\delta_{14} \mid others; data) = (\delta_{14} \mid \mathbf{u}, \boldsymbol{\theta_1}, \mathbf{Z}_1)$ is

$$IG\left(a_{14} + \frac{I}{2}, b_{14} + \frac{1}{2}\sum_{i=1}^{I}(u_{i4} - \theta_{13} - z_{1i})^2\right).$$

(ix) $(\delta_{2j} \mid others; data) = (\delta_{2j} \mid \mathbf{v}, \boldsymbol{\theta_2}, \mathbf{Z}_2)$ for $j = 1, \ldots, 3$ is

$$IG\left(a_{2j} + \frac{I}{2}, b_{2j} + \frac{1}{2}\sum_{i=1}^{I}(v_{ij} - \theta_{2j} - z_{2i})^2\right).$$

(x) $[\delta_1 \mid others; data] = [\delta_1 \mid \mathbf{Z}_1, \mathbf{Z}_2, \rho_1, \rho_2, \rho_3, \delta_2]$ is proportional to

$$\frac{\exp\left\{\frac{-\mathbf{Z}_1^T\mathbf{B}_1\mathbf{Z}_1}{2\delta_1(1-\rho_3^2)} + \frac{\rho_3}{(1-\rho_3^2)}\frac{\mathbf{Z}_1^T\mathbf{B}_1^{\frac{1}{2}}\mathbf{B}_2^{\frac{1}{2}}\mathbf{Z}_2}{\sqrt{\delta_1\delta_2}} - \frac{b_{z1}}{\delta_1}\right\}}{\delta_1^{a_{z1}+\frac{I}{2}+1}}.$$

Or equivalently $[\delta_1^* \mid others; data] = [\delta_1^* \mid \mathbf{Z}_1, \mathbf{Z}_2, \rho_1, \rho_2, \rho_3, \delta_2^*]$ is

$$\propto \exp\left\{k_{11}\ln(\delta_1^*) + k_{12}\delta_1^* - k_{13}\delta_1^{*2}\right\},$$

where $\delta_1^* = \delta_1^{-1/2}$, $\delta_2^* = \delta_2^{-1/2}$, $k_{11} = 2a_{z1} + I - 1$,

$$k_{12} = \frac{\rho_3\delta_2^*}{1 - \rho_3^2}\mathbf{Z}_1^T\mathbf{B}_1^{\frac{1}{2}}\mathbf{B}_2^{\frac{1}{2}}\mathbf{Z}_2, \text{ and } k_{13} = \frac{\mathbf{Z}_1^T\mathbf{B}_1\mathbf{Z}_1}{2(1 - \rho_3^2)} + b_{z1}.$$

(xi) $[\delta_2 \mid others; data] = [\delta_2 \mid \mathbf{Z}_1, \mathbf{Z}_2, \rho_1, \rho_2, \rho_3, \delta_1]$ is proportional to:

$$\frac{\exp\left\{\frac{-\mathbf{Z}_2^T\mathbf{B}_2\mathbf{Z}_2}{2\delta_2(1-\rho_3^2)} + \frac{\rho_3}{(1-\rho_3^2)}\frac{\mathbf{Z}_2^T\mathbf{B}_2^{\frac{1}{2}}\mathbf{B}_1^{\frac{1}{2}}\mathbf{Z}_1}{\sqrt{\delta_1\delta_2}} - \frac{b_{z2}}{\delta_2}\right\}}{\delta_2^{a_{z2}+\frac{I}{2}+1}}.$$

It is equivalent to that $[\delta_2^* \mid others; data] = [\delta_2^* \mid \mathbf{Z}_1, \mathbf{Z}_2, \rho_1, \rho_2, \rho_3, \delta_1^*]$ is

$$\propto \exp\left\{k_{21}\ln(\delta_2^*) + k_{22}\delta_2^* - k_{23}\delta_2^{*2}\right\},$$

where $\delta_1^* = \delta_1^{-1/2}$, $\delta_2^* = \delta_2^{-1/2}$, $k_{21} = 2a_{z2} + I - 1$,

$$k_{22} = \frac{\rho_3\delta_1^*}{1 - \rho_3^2}\mathbf{Z}_2^T\mathbf{B}_2^{\frac{1}{2}}\mathbf{B}_1^{\frac{1}{2}}\mathbf{Z}_1, \text{ and } k_{23} = \frac{\mathbf{Z}_2^T\mathbf{B}_2\mathbf{Z}_2}{2(1 - \rho_3^2)} + b_{z2}.$$

(xii) $[\rho_1 \mid others; data] = [\rho_1 \mid \mathbf{Z}_1, \mathbf{Z}_2, \rho_2, \rho_3, \delta_1, \delta_2]$ is proportional to

$$|\mathbf{I} - \rho_1\mathbf{C}|^{\frac{1}{2}}$$
$$\times \exp\left\{\frac{-\mathbf{Z}_1^T(\mathbf{I} - \rho_1\mathbf{C})\mathbf{Z}_1}{2(1 - \rho_3^2)\delta_1}\right.$$
$$\left. + \frac{\rho_3}{(1 - \rho_3^2)}\frac{\mathbf{Z}_1^T(\mathbf{I} - \rho_1\mathbf{C})^{\frac{1}{2}}\mathbf{B}_2^{\frac{1}{2}}\mathbf{Z}_2}{\sqrt{\delta_1\delta_2}}\right\}.$$

(xiii) $[\rho_2 \mid others; data] = [\rho_2 \mid \mathbf{Z}_1, \mathbf{Z}_2, \rho_1, \rho_3, \delta_1, \delta_2]$ is proportional to

$$|\mathbf{I} - \rho_2\mathbf{C}|^{\frac{1}{2}}$$
$$\times \exp\left\{\frac{-\mathbf{Z}_2^T(\mathbf{I} - \rho_2\mathbf{C})\mathbf{Z}_2}{2(1 - \rho_3^2)\delta_2}\right.$$
$$\left. + \frac{\rho_3}{(1 - \rho_3^2)}\frac{\mathbf{Z}_2^T(\mathbf{I} - \rho_2\mathbf{C})^{\frac{1}{2}}\mathbf{B}_1^{\frac{1}{2}}\mathbf{Z}_1}{\sqrt{\delta_2\delta_1}}\right\}.$$

(xiv) $[\rho_3 \mid others; data] = [\rho_3 \mid \mathbf{Z}_1, \mathbf{Z}_2, \rho_1, \rho_2, \delta_1, \delta_2]$ is proportional to

$$\frac{|\mathbf{B_1}|^{\frac{1}{2}}|\mathbf{B_2}|^{\frac{1}{2}}}{[(1 - \rho_3^2)\delta_1\delta_2]^{\frac{I}{2}}}$$
$$\times \exp\left\{-\frac{1}{2}\left[\frac{\mathbf{Z}_1^T\mathbf{B}_1\mathbf{Z}_1}{(1 - \rho_3^2)\delta_1} - 2\frac{\rho_3\mathbf{Z}_1^T\mathbf{B}_1^{\frac{1}{2}}\mathbf{B}_2^{\frac{1}{2}}\mathbf{Z}_2}{(1 - \rho_3^2)\sqrt{\delta_1\delta_2}}\right.\right.$$
$$\left.\left. + \frac{\mathbf{Z}_2^T\mathbf{B}_2\mathbf{Z}_2}{(1 - \rho_3^2)\delta_2}\right]\right\}.$$

Proposition 2.

(a) *The conditional densities of u_{ij} and v_{ij} in Proposition 1 are both log concave.*

(b) *The conditional densities of δ_1^* and δ_2^* in Proposition 1 are both log concave.*

Proof. We only give a proof for part (b). Let $f(\delta_1^*)$ be the conditional density of δ_1^*. Clearly, up to a positive constant,

$$f(\delta_1^*) = \exp\left\{ k_{11}\ln(\delta_1^*) + k_{12}\delta_1^* - k_{13}\delta_1^{*2} \right\},$$

where k_{11}, k_{12} and k_{13} are defined in Proposition 1. Thus,

$$
\begin{aligned}
\log f(\delta_1^*) &= k_{11}\ln(\delta_1^*) + k_{12}\delta_1^* - k_{13}\delta_1^{*2}, \\
\frac{\partial^2 \log(f(\delta_1^*))}{\partial \delta_1^{*2}} &= -\frac{k_{11}}{\delta_1^{*2}} - 2k_{13}.
\end{aligned}
$$

The second derivative of $\log f(\delta_1^*)$ is negative since $k_{11} > 0$ and $k_{13} > 0$. The log-concavity of density of δ_2^* is similar and is omitted. $\square$

Received 31 October 2013

REFERENCES

ALBERT, J. H. and GUPTA, A. K. (1985). Bayesian methods for binomial data with applications to a nonresponse problem, *Journal of the American Statistical Association* **80**: 167–174. MR0786603

BASU, D. and DE BRAGANCA PEREIRA, C. A. (1982). On the Bayesian analysis of categorical data: The problem of nonresponse, *Journal of Statistical Planning and Inference* **6**: 345–362. MR0667914

BESAG, J. (1974). Spatial interaction and the statistical analysis of lattice systems (with discussion), *Journal of the Royal Statistical Society, Series B: Methodological* **36**: 192–236. MR0373208

CARLIN, B. P. and LOUIS, T. A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*, Chapman & Hall Ltd.

CHEN, M.-H., DEY, D. K. and SHAO, Q.-M. (1999). A new skewed link model for dichotomous quantal response data, *Journal of the American Statistical Association* **94**(448): 1172–1186. MR1731481

CLAYTON, D. and KALDOR, J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping, *Biometrics* **43**: 671–681.

DE LEEUW, E. and DE HEER, W. (2002). Trends in household survey nonresponse: A longitudinal and international comparison, *Survey Nonresponse,* pp. 41–54, eds. R. M. Groves, D. A. Dillmam, J. L. Eltinge, and R. J. A. Little, John Wiley and Sons, New York.

DILLMAN, D. A. (2000). *Mail and Internet Surveys: The Tailored Design Method*, John Wiley & Sons.

FILION, F. L. (1980). *Human Surveys in Wildlife Management*, Wildlife Management Techniques Manual, The Wildlife Society, Washington DC, pp. 441–453.

FORSTER, J. J. and SMITH, P. W. F. (1998). Model-based inference for categorical survey data subject to non-ignorable non-response (Disc: P89-102), *Journal of the Royal Statistical Society, Series B: Statistical Methodology* **60**: 57–70. MR1625664

GAO, X., HE, C. and SUN, D. (2014). A Bayesian spatial model with auxiliary covariates to assess and adjust nonignorable nonresponse, *Spatial Statistics* **8**: 122–144. Spatial Statistics Miami.

GELFAND, A. E. and SMITH, A. F. M. (1990). Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association* **85**: 398–409. MR1141740

GELMAN, A., CARLIN, J. B., STERN, H. S. and RUBIN, D. B. N. E. (2004). *Bayesian Data Analysis*, Chapman and Hall/CRC. MR2027492

GILKS, W. R. E., RICHARDSON, S. E. and SPIEGELHALTER, D. J. E. (1996). *Markov Chain Monte Carlo in Practice*, Chapman & Hall Ltd. MR1397966

GILKS, W. R. and WILD, P. (1992). Adaptive rejection sampling for Gibbs sampling, *Applied Statistics* **41**: 337–348.

HE, Z. and SUN, D. (2000). Hierarchical Bayes estimation of hunting success rates with spatial correlations, *Biometrics* **56**(2): 360–367.

JOHNSON, R. A. and WICHERN, D. W. (1988). *Applied Multivariate Statistical Analysis*, Prentice-Hall Inc.

LESSLER, J. T. and KALSBEEK, W. D. (1992). *Nonsampling Error in Surveys*, John Wiley & Sons. MR1193229

LITTLE, R. J. A. and RUBIN, D. B. (2002). *Statistical Analysis with Missing Data*, John Wiley & Sons. MR1925014

NANDRAM, B. and CHOI, J. W. (2002a). A Bayesian analysis of a proportion under non-ignorable non-response, *Statistics in Medicine* **21**(9): 1189–1212.

NANDRAM, B. and CHOI, J. W. (2002b). Hierarchical Bayesian nonresponse models for binary data from small areas with uncertainty about ignorability, *Journal of the American Statistical Association* **97**(458): 381–388. MR1941460

NANDRAM, B., COX, L. H. and CHOI, J. W. (2005). Bayesian analysis of nonignorable missing categorical data: An application to bone mineral density and family income, *Survey Methodology* **31**(2): 213–225.

OLESON, J. J. and HE, C. Z. (2008). Adjusting nonresponse bias at subdomain levels using multiple response phases, *Biometrical Journal* **50**(1): 58–70. MR2414638

RAO, J. N. K. (2003). *Small Area Estimation*, John Wiley & Sons. MR1953089

ROBERT, C. P. and CASELLA, G. (2004). *Monte Carlo Statistical Methods*, Springer-Verlag Inc. MR2080278

STASNY, E. A. (1991). Hierarchical models for the probabilities of a survey classification and nonresponse: An example from the National Crime Survey, *Journal of the American Statistical Association* **86**: 296–303.

WAKEFIELD, J. C., GELFAND, A. E. and SMITH, A. F. M. (1991). Efficient generation of random variates via the ratio-of-uniforms method, *Statistics and Computing* **1**: 129–133.

WOODARD, R. (1999). *Bayesian Hierarchical Models for Hunting Success Rates*, PhD thesis, University of Missouri-Columbia. MR2700015

Xiaoming Gao
Central Regional Office and Conservation Research Center
Missouri Department of Conservation
3500 East Gans Road
Columbia, MO 65201
USA
E-mail address: Sherry.Gao@mdc.mo.gov

Chong He
Department of Statistics
University of Missouri
Columbia, MO 65211
USA
E-mail address: hezh@missouri.edu

Dongchu Sun
Department of Statistics
University of Missouri
Columbia, MO 65211
USA
E-mail address: SunD@missouri.edu