Bayesian Analysis of a Self-selection Model with Multiple Outcomes

Jieyu Gao

February, 2022

1 Introduction

This paper proposed a parametric self-selection Bayesian model with one binary treatment and two outcome variables. One outcome variable is a count variable and the other one is a continuous variable. Allowing for endogenous selection, we are interested in studying the treatment impact on the conditional mean of the outcome variables. We applied the model to two datasets to study the impact of private insurance on healthcare expenditures and number of Physician office visits.

Sample selection occurs when the sample observed is not a random sample and thus can not represent the population of interest. Heckman (1979) pointed out that ignoring the sample selection would usually cause an inconsistent and biased estimators and introduced Heckman correction to address this issue. In the treatment model, we only observe the treated outcome for the individuals in the treated group, and untreated outcome for the individuals in the untreated group, and the sample selection bias may occur if the selection decisions are not random. Chib et al. (2009) developed a Bayesian model to analyze the data with sample selection and endogeneity problem.

Many papers discussed possible selection concerns in health insurance industry. Cutler and Zeckhauser (1998) discussed some evidence of adverse selection. Keane and Stavrunova (2016) proposed a model to estimate the moral hazard and selection in the Medigap market. They found evidence that people with better health conditions are more likely to purchase the Medigap. Sapelli and Vial (2003) studied the self-selection using the Chilean physician visits data and hospital days data. They found some evidence supporting the existence of self-selection. Shane (2012) found evidence and examines the sources of selection on unobservables. Nghiem and Graves (2019) estimates the impact of moral hazard and selection bias using a panel data from the private insurance industry in Australia and found evidence of advantageous selection that risk adverse people are more likely to purchase the health insurance.

On the other hand, Reschovsky et al. (2000) suggested that there is no substantial difference in terms of hospital use across different insurance plan types. Cardon and Hendel (2001) found that after controlling for many observables, there is no evidence of asymmetric information that has impact on the insurance plan choices. Munkin and Trivedi (2003) found weak evidence of self-selection in his model estimating the impact of health insurance on health care utilization.

We extend the method used in Munkin and Trivedi (2003) to simplify the computation process by assuming that the continuous variable is of a log-normal distribution, instead of assuming an exponential distribution where the parameter is log-normally distributed. We applied the method to two datasets to study the impact of public or private health insurance on health care expenditure and number of doctor's office Visit. The relevant data used in this application is 1996 MEPS data and the 1987 NMES data. We also estimated the data using four more parsimonious models. The marginal likelihood for each of the models is calculated for comparison.

Developed on the basis of many recent Bayesian research of Albert and Chib (1993), Chib et al. (1998) and Chib et al. (2009), we use the Markov chain Monte Carlo (MCMC) method to estimate the parameters.

The rest of the paper is organized as follows. Section 2 specifies the model. The estimation method is presented in Section 3. Section 4 presents the simulation results. Section 5 presents the empirical results and concludes the paper.

2 Model

2.1 Likelihood

The treatment variable is denoted by T_i, where $T_i = 1$ if subject i is treated, and $T_i = 0$ if subject i is in the control group. We observe N independent observations and use y_{1ti} and y_{it2} to denote the outcome variables, where y_{it1} is the discrete potential outcome variable and y_{it2} is the potential continuous outcome variable, $t \in \{0, 1\}$ to denote the treatment status. Let X_1 denote the exogenous covariates matrix for y_1 and X_2 denote the exogenous covariates matrix for y_2. We assume that

$$y_{i1k} \sim \text{Poisson}(\mu_{ik}) \tag{1}$$

Let's use d_i to denote the treatment variable for subject i.

For each subject i in the sample, the observed response is:

$$y_{i1} = y_{i10} + (y_{i11} - y_{i10})T_i \tag{2}$$

$$y_{i2} = y_{i20} + (y_{i21} - y_{i20})T_i \tag{3}$$

z_i is a latent variable that relates to T_i, and $T_i = \mathbb{1}\{z_i \geq 0\}$.

Thus we are going to either observe the y_{i11} and y_{i21} if the subject i is in the treatment group, and y_{i10} and y_{i20} if the subject i is untreated.

The model can be represented represented as following equation:

$$g_i = X_i\beta + \epsilon_i \tag{4}$$

where

$$g_i = (\ln(\mu_{i0}), y_{i20}, \ln(\mu_{i1}), y_{i21}, z_i)', \quad X_i = \begin{pmatrix} x'_{i10} & 0 & 0 & 0 & 0 \\ 0 & x'_{i20} & 0 & 0 & 0 \\ 0 & 0 & x'_{i11} & 0 & 0 \\ 0 & 0 & 0 & x'_{i21} & 0 \\ 0 & 0 & 0 & 0 & x'_{i5} \end{pmatrix}$$

$\beta = (\beta'_{10}, \beta'_{20}, \beta'_{11}, \beta'_{21}, \beta'_5)'$ and $\epsilon_i = (\epsilon_{i1}, \epsilon_{i2}, \epsilon_{i3}, \epsilon_{i4}, \epsilon_{i5})'$.

If $z_i > 0$, we observe $(\ln(\mu_{i1}), y_{i21})$ and otherwise, we observe $(\ln(\mu_{i0}), y_{i20})$. We define the vectors $s_{i0} = (\ln(\mu_{i0}), y_{i20}, z_i)'$ and $s_{i1} = (\ln(\mu_{i1}), y_{i21}, z_i)'$. We also let $N_1 = \{i : T_i = 1\}$ and $N_0 = \{i : T_i = 0\}$ to denote the treated and untreated samples correspondingly. Additionally, we assume that

$$X_{i1} = \begin{pmatrix} x'_{i11} & 0 & 0 \\ 0 & x'_{i21} & 0 \\ 0 & 0 & x_{i5} \end{pmatrix} \text{ and } X_{i0} = \begin{pmatrix} x'_{i10} & 0 & 0 \\ 0 & x'_{i20} & 0 \\ 0 & 0 & x_{i5} \end{pmatrix}$$

Assume that $\epsilon_i \sim N(0, \Omega)$. Let's define:

$$\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} & \omega_{13} & \omega_{14} & \omega_{15} \\ \omega_{21} & \omega_{22} & \omega_{23} & \omega_{24} & \omega_{25} \\ \omega_{31} & \omega_{32} & \omega_{33} & \omega_{34} & \omega_{35} \\ \omega_{41} & \omega_{42} & \omega_{43} & \omega_{44} & \omega_{45} \\ \omega_{51} & \omega_{52} & \omega_{53} & \omega_{54} & 1 \end{pmatrix}$$

Due to the missing outcomes, we can't identify Ω_{13}, Ω_{14}, Ω_{23} and Ω_{24}. We also assume that $\Omega_{55} = 1$.

Thus the covariance matrix that can be identified is:

$$\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} & . & . & \omega_{15} \\ \omega_{21} & \omega_{22} & . & . & \omega_{25} \\ . & . & \omega_{33} & \omega_{34} & \omega_{35} \\ . & . & \omega_{43} & \omega_{44} & \omega_{45} \\ \omega_{51} & \omega_{52} & \omega_{53} & \omega_{54} & 1 \end{pmatrix}$$

We define:

$$\Omega_0 = \begin{pmatrix} \omega_{11} & \omega_{12} & \omega_{15} \\ \omega_{21} & \omega_{22} & \omega_{25} \\ \omega_{51} & \omega_{52} & 1 \end{pmatrix}, \quad \Omega_1 = \begin{pmatrix} \omega_{33} & \omega_{34} & \omega_{35} \\ \omega_{43} & \omega_{44} & \omega_{45} \\ \omega_{53} & \omega_{54} & 1 \end{pmatrix}$$

$$J_0 = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I} \end{pmatrix}, \quad J_1 = \begin{pmatrix} 0 & 0 & \mathbf{I} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I} \end{pmatrix}$$

Thus we have $J_0\beta = (\beta'_{10}, \beta'_{20}, \beta'_5)$, and $J_1\beta = (\beta'_{11}, \beta'_{21}, \beta'_5)$. Under the assumption that $\omega_{55} = 1$, we can estimate Ω_0 and Ω_1 separately.

Thus for $i \in N_0$, we have:

$$s_{i0} \propto |\Omega_0|^{-1/2} \exp\left\{ -\frac{1}{2}(s_{i0} - X_{i0}J_0\beta)'\Omega_0^{-1}(s_{i0} - X_{i0}J_0\beta) \right\} \tag{5}$$

, and for $i \in N_1$, we have:

$$s_{i1} \propto |\Omega_1|^{-1/2} \exp\left\{ -\frac{1}{2}(s_{i1} - X_{i1}J_1\beta)'\Omega_1^{-1}(s_{i1} - X_{i1}J_1\beta) \right\} \tag{6}$$

Since we don't observe the data $\ln(\mu_i)$ directly from the data, the complete data density function is thus:

$$\begin{aligned} f(y_1, y_2, z, \ln(\mu)|\theta) &= p(\ln(\mu), y_2, z|\theta)p(y_1|\ln(\mu)) \\ &= \left[\prod_{i \in N_1} f(s_{i1}|\theta) f_{Po}(y_{1i}|\ln(\mu_{i1}))\mathbb{1}\{z_i \geq 0\}\right]\left[\prod_{i \in N_0} f(s_{i0}|\theta) f_{Po}(y_{1i}|\ln(\mu_{i0}))\mathbb{1}\{z_i < 0\}\right] \end{aligned} \tag{7}$$

Under the assumption that $\omega_{55} = 1$, we can estimate the Ω_0 and Ω_1 separately.

Let's partition Ω_t, where $t \in \{0, 1\}$ as follows:

$$\Omega_t = \begin{pmatrix} \underset{(2\times 2)}{\Omega_{11}^t} & \Omega_{21}^t \\ \Omega_{12}^t & 1 \end{pmatrix} = \begin{pmatrix} \Omega_{11}^{\tilde{t}} & \Omega_{12}^{\tilde{t}} \\ \Omega_{21}^{\tilde{t}} & \underset{(2\times 2)}{\Omega_{22}^{\tilde{t}}} \end{pmatrix}$$

To simplify the notation, we define the following functions:

$$\Omega_{11\cdot 2}^t = \Omega_{11}^t - \Omega_{21}^t \Omega_{12}^t$$

2.2 Prior Distribution

We assume that β has a jointly normal distribution with mean b_0 and covariance matrix B_0.

Thus for the prior density function of the parameters β is as follows:

$$\pi(\beta) = f_N(\beta|b_0, B_0) \tag{8}$$

We assume that $\Omega_{11\cdot 2}^t \sim IW(\nu = 5, D_t)$ for t in $\{0, 1\}$.

Then the conditional distribution of $\Omega_{12}^t|\Omega_{11\cdot 2}^t \sim MN(\lambda_t, \Omega_{11\cdot 2}^t)$.

3 Estimation

3.1 Sampling β

The posterior distribution for β is $N(b, B)$, where:

$$b = B\left(B_0^{-1}b_0 + \sum_{i\in N_0} J_0' X_{i0}' \Omega_0^{-1} y_{i0} + \sum_{i\in N_1} J_1' X_{i1}' \Omega_1^{-1} y_{i1}\right)$$

, and

$$B = \left(B_0^{-1} + \sum_{i\in N_0} J_0' X_{i0}' \Omega_0^{-1} X_{i0} J_0 + \sum_{i\in N_1} J_1' X_{i1}' \Omega_1^{-1} X_{i1} J_1\right)^{-1}.$$

3.2 Sampling z_i

For $t \in \{0, 1\}$, given $\ln(\mu_{it})$, $\ln(y_{i2t})$, d, β and Ω, z_i is of a truncated normal distribution:

$$\text{For } i \in N_t: \quad (z_i | \ln(\mu_{it}), y_{i2t}, \beta, \Omega_t) \sim TN(\mu_{zt}, \Omega_{zt}) \tag{9}$$

where $z_i \in (-\infty, 0)$ if $i \in N_0$ and $z_i \in [0, +\infty)$ if $i \in N_1$. Please note that $\mu_{zt} = x_{i5}'\beta_5 + \Omega_{12}^t(\Omega_{11}^t)^{-1}\left(\begin{pmatrix} \ln(\mu_{it}) - x_{it}'\beta_{1t} \\ y_{i2t} - x_{i2}'\beta_{2t} \end{pmatrix}\right)$, $\Omega_{zt} = 1 - \Omega_{12}^t(\Omega_{11}^t)^{-1}\Omega_{21}^t$.

3.3 Sampling $\ln(\mu_{it})$

For $i \in N_t$, the conditional distribution of $\ln(\mu_{it})$ given y_{i1t}, y_{i2t} and z_i is

$$(\ln(\mu_{it}) | y_{i2t}, z_i, \beta, \Omega) \sim N(\mu_{1|2}^t, \sigma_{1|2}^t)$$

where $\mu_{1|2}^t = x_{i1t}'\beta_{1t} + \tilde{\Omega}_{12}^t(\tilde{\Omega}_{22}^t)^{-1}\begin{pmatrix} y_{i2t} - x_{i2t}'\beta_{2t} \\ z_i - x_{i5}'\beta_5 \end{pmatrix}$, and $\sigma_{1|2}^t = \tilde{\Omega}_{11}^t - \tilde{\Omega}_{12}^t(\tilde{\Omega}_{22}^t)^{-1}\tilde{\Omega}_{21}^t$. Thus the posterior distribution for $\ln(\mu_{it})$ is as follows:

$$\pi(\ln(\mu_{it}) | y_{i1t}, y_{i2t}, z_i, \beta, \Omega_t) = f_N(\mu_{1|2}^t, \sigma_{1|2}^t)\pi(y_{i1t} | \ln(\mu_{i1t})) = f_N(\mu_{1|2}^t, \sigma_{1|2}^t)\frac{\mu^{y_{i1t}} e^{-\mu_{it}}}{y_{i1t}!} \tag{10}$$

We use metropolis-hasting with random walk to sample the $\ln(\mu_{it})$. We use m_i^* to denote the proposed $\ln(\mu_{it})$ and m_i to denote the current $\ln(\mu_{it})$. The proposed density is denoted as $q(m_i, m_i^*) = \phi(m_i^* | m_i, \tau((\sigma_{1|2}^t)^{-1} + y_{i1t}^{-1})^{-1})$. The acceptance rate is thus defined as:

$\alpha(m_i, m_i^*) = \min\{\frac{\pi(m_i^*|y_{i1t},y_{i2t},\beta,z_i,\Omega)q(m_i,m_i^*)}{\pi(m_i|y_{i1t},y_{i2t},\beta,z_i,\Omega)q(m_i^*,m)}, 1\} = \min\{\frac{\pi(m_i^*|y_{i1t},y_{i2t},\beta,z_i,\Omega)}{\pi(m_i|y_{i1t},y_{i2t},\beta,z_i,\Omega)}, 1\}.$

We draw a random number from $p \in U(0,1)$ and if $p \leq \alpha$, the proposed value is accepted, and otherwise, the proposed value is rejected.

3.4 Sampling Ω_{21}^t

Note that for $i \in N_t$:

$$\pi(\Omega_{21}^t|\Omega_{11\cdot 2}^t, \beta, \ln(\mu_{it}), \ln(y_{i2t}), z) = \pi(\Omega_{21}^t|\Omega_{11\cdot 2}^t)\pi\left(\begin{pmatrix}\ln(\mu_{it}) \\ y_{i2t}\end{pmatrix}|z\right)$$

$$= f_N(\lambda_t, \Omega_{11\cdot 2}^t)\prod_{i\in N_t} f_N((\ln(\mu_{it}), y_{i2t})'|\mu_{2|3}^t, \Omega_{11\cdot 2}^t)) \tag{11}$$

where $\mu_{2|3}^t = \begin{pmatrix}x_{i1t}'\beta_{1t} \\ x_{i2t}'\beta_{2t}\end{pmatrix} + \Omega_{21}^t(z_i - x_{i5}'\beta_5)$.

Thus the posterior distribution for Ω_{21}^t is also a normal distribution.

The posterior distribution: $\Omega_{21}^t|\Omega_{11\cdot 2}^t, \beta \sim N([\epsilon_{zt}\epsilon_{zt}(\Omega_{11\cdot 2}^t)^{-1} + (\Omega_{11\cdot 2}^t)^{-1}]^{-1}[(\Omega_{11\cdot 2}^t)^{-1}\epsilon_t'\epsilon_{zt} + (\Omega_{11\cdot 2}^t)^{-1}\lambda_t], [\epsilon_{zt}'\epsilon_{zt}(\Omega_{11\cdot 2}^t)^{-1} + (\Omega_{11\cdot 2}^t)^{-1}]^{-1})$, where $\epsilon_{zt} = z - x_5\beta_5$ for all observations in group N_t, and $\epsilon_t = \Big(\ln(\mu_t) - x_{1t}\beta_{1t}, \quad y_{2t} - x_{2t}\beta_{2t}\Big)$.

3.5 Sampling $\Omega_{11\cdot 2}^t$

For $i \in N_t$:

$$\pi(\Omega_{11\cdot 2}^t|\Omega_{21}^t, \beta, \ln(\mu_{it}), y_{i2t}, z_i) = \pi(\Omega_{11\cdot 2}^t)\pi(\Omega_{21}^t|\Omega_{11\cdot 2}^t)\pi\left(\begin{pmatrix}\ln(\mu_{it}) \\ \ln(y_{i2t})\end{pmatrix}|z_i\right)$$

$$= \pi(\Omega_{11\cdot 2}^t)f_N(\lambda_t, \Omega_{11\cdot 2}^t)\prod_{i\in N_t} f_N((\ln(\mu_{it}), y_{i2t})'|\mu_{2|3}^t, \Omega_{11\cdot 2}^t))$$

$$\propto |\Omega_{11\cdot 2}^t|^{\frac{\nu+4+N_t}{2}}\exp[-\frac{1}{2}(tr\left(\left[D_t + (\Omega_{21}^t - \lambda_t)(\Omega_{21}^t - \lambda_t)' + \sum_{i\in N_t}(\epsilon_{it} - \Omega_{21}^t\epsilon_{izt})(\epsilon_{it} - \Omega_{12}^t\epsilon_{izt})'\right](\Omega_{11\cdot 2}^t)^{-1}\right)]$$

The posterior distribution for $\Omega_{11\cdot 2}^t$ is as follows:

$$\Omega_{11\cdot 2}^t \sim \text{Inverse Wishart}(\nu+1+N_t, D_t + (\Omega_{21}^t - \lambda_t)(\Omega_{21}^t - \lambda_t)' + \sum_{i\in N_t}(\epsilon_{it} - \Omega_{21}^t\epsilon_{izt})(\epsilon_{it}^A - \Omega_{21}^t\epsilon_{izt})') \tag{12}$$

3.6 Model Comparison

We decided to compare the following models: M_0, the baseline model with no restrictions; model M_1, the baseline model with constraints $\Omega_{12}^0 = \Omega_{12}^1 = (0,0)$; M_2, the model with constant treatment effect, and M_3, model M_2 with a restriction that $\Omega_{12} = (0,0)$. Notice that M_2 is a parsimonious version of M_0 with the assumption that the slope for the covariates in the medical utilization equations are the same for the treated group and control group. The model specification and estimation for M_2 is discussed in Appendix A.

Because the marginal likelihood in model M_i can be written as:

$$m(y|M_i) = \frac{f(y|M_i, \theta_i)\pi(\theta_i|M_i)}{\pi(\theta_i|y, M_i)} \tag{13}$$

According to Chib (1995), we can estimate the marginal likelihood by MCMC methods at an appropriate point θ^*, and by taking logarithms on both side of the equation, we can obtain:

$$\log \hat{m}(y|M_i) = \log f(y|M_i, \theta^*) + \log \pi(\theta^*|M_i) - \log \hat{\pi}(\theta^*|y, M_i) \tag{14}$$

The first two terms can be calculated easily. Suppose that $\theta^* = (\theta_1^*, ..., \theta_B^*)$. Let's use $\{\theta^{(g)}\} = \{\theta^{(1)}, ..., \theta^{(G)}\}$ be G draws from the posterior distribution $\pi(\theta|y, M_i)$ for model M_i using the MCMC method.

Let's use z to denote the latent variables.

Based on Jeliazkov and Lee (2010) and Ritter and Tanner (1992),

$$\pi(\theta^*|y) = \int K(\theta, \theta^*|y, z)\pi(\theta, z|y)d\theta dz \tag{15}$$

where $K(\theta, \theta^*|y, z) = \prod_{r=1}^{B} \pi(\theta_r^*|y, \{\theta_s^*\}(s < r), \{\theta_s^{(g)}\}(s > r), \{z^{(g)}\})$.

The estimator is

$$\hat{\pi}(\theta^*|y) = G^{-1} \sum_{g=1}^{G} K(\theta^{(g)}, \theta^*|y_i, \{z^{(g)}\}) \tag{16}$$

In our model, $\theta = (\beta, \Omega_{12}^0, \Omega_{12}^1, \Omega_{11\cdot 2}^0, \Omega_{11\cdot 2}^1)'$, and the latent variables are z, $\ln(\mu_0)$ and $\ln(\mu_1)$. We choose θ^* as the mean of the the first 5000 iterations, and calculate $\hat{\pi}(\theta^*|y)$ using the remaining iterations.

The likelihood function can be numerically calculated using either Gauss–Legendre quadrature or importance sampling. We use student-t distribution with degree of freedom 5 as the proxy distribution for the importance sampling estimation process.

4 Simulation

We randomly generated 10,000 observations from the unrestricted model and estimate the model using the method introduced in the previous session.

4.1 Data Generating Process

$$\Omega = \begin{pmatrix} 2.3 & 0.9 & 0.8 & 1.1 & 0.5 \\ 0.9 & 2 & 1 & 0.8 & 0.7 \\ 0.8 & 1 & 1.8 & 0.6 & 0.6 \\ 1.1 & 0.8 & 0.6 & 2.5 & 0.7 \\ 0.5 & 0.7 & 0.6 & 0.7 & 1 \end{pmatrix}$$

$\beta_{10} = (-1, 1)'$, $\beta_{20} = (-2, 2)'$, $\beta_{11} = (1, -1)'$, $\beta_{21} = (-1, -1)'$ and $\beta_5 = (-0.5, -1, 1)'$. The covariates are $x_{10} = [1, \nu_1]$, $x_{20} = [1, \nu_1]$, $x_{11} = [1, \nu_1]$, $x_{21} = [1, \nu_1]$ and $x_5 = [1, \nu_1, \nu_2]$, where $\nu_i \sim N(0, 4)$ for $j \in \{1, 2\}$. We randomly sample the $(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_5)$ from the distribution $N(0, \Omega)$. $\ln(\mu_0)$ and $\ln(\mu_1)$ are calculated correspondingly according to the original model, and y_1 and y_2 are sampled based on the original model.

4.2 Simulation Results

We estimated the generated data using 50,000 iterations. The estimated result can be shown in Table 1 and Table 2. Figure 1 shows the posterior distributions for the first element of β_{10} and ω_{13}. The estimation results using the artificial data show the estimated values are centered at values that are close to the true parameter values. The 95% credible intervals are obtained from the converged empirical estimate distribution by obtaining the 2.5% and 97.5% quantile.

The log marginal likelihood are summarized in Table 3. We used the method described in section 3.6 to estimate the log marginal likelihood for all four models. We used both the Gauss–Legendre quadrature and importance sampling to estimate the likelihood and calculated the corresponding marginal likelihood. We obtain very similar results using those two different methods. The results suggest an extremely strong evidence for model M_0 compared to all the other three models.

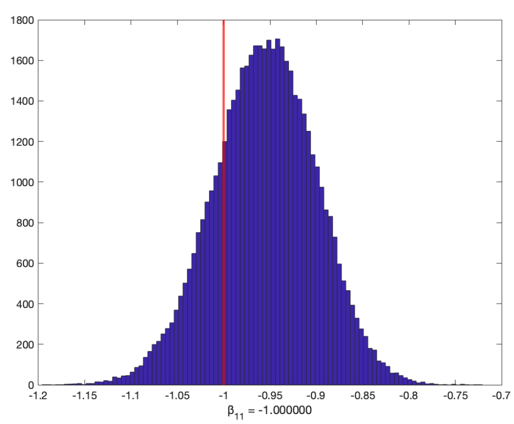
Table 1

M_0 and M_1: Estimation for generated data

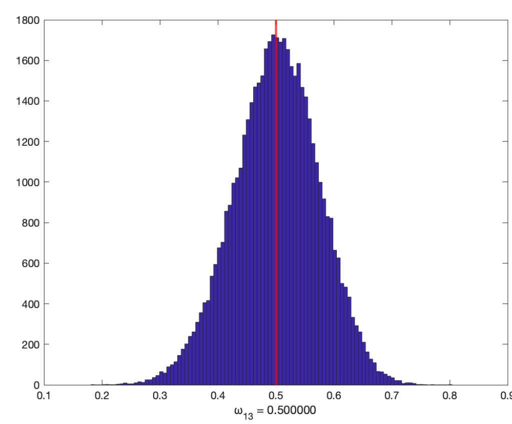
	True Value	M_0 Unrestricted			M_1 Restricted		
		Mean	SD	95% Credible Interval	Mean	SD	95% Credible Interval
CONST1	-1	-0.9555	0.0554	(-1.0673, -0.8500)	-1.1656	0.0486	(-1.2626, -1.0717)
v1	1	0.9829	0.0203	(0.9439 , 1.0230)	1.0345	0.0194	(0.9968 , 1.0729)
CONST2	-2	-1.9515	0.0330	(-2.0159 , -1.8862)	-2.2664	0.0261	(-2.3177, -2.2154)
v1	2	1.9805	0.0136	(1.9538, 2.0071)	2.0586	0.0126	(2.0339, 2.0833)
CONST3	1	1.0141	0.0292	(0.9561, 1.0708)	1.2014	0.0245	(1.1532, 1.2494)
v1	-1	-1.0001	0.0127	(-1.0249 , -0.9753)	-0.9480	0.0120	(-0.9717, -0.9245)
CONST4	-1	-0.9705	0.0285	(-1.0267, -0.9148)	-0.7576	0.0236	(-0.8042, -0.7117)
v1	-1	-0.9902	0.0129	(-1.0154, -0.9651)	-0.9291	0.0121	(-0.9530, -0.9053)
CONST5	0.5	0.5029	0.0227	(0.4589 , 0.5478)	0.4992	0.0227	(0.4548, 0.5432)
v1	-1	-0.9987	0.0205	(-1.0389, -0.9583)	-0.9922	0.0212	(-1.0343, -0.9508)
v2	1	1.0074	0.0207	(0.9671 , 1.0481)	1.0052	0.0212	(0.9630 , 1.0464)
w11	2.3	2.3953	0.0799	(2.2425 , 2.5568)	2.3663	0.0803	(2.2146 , 2.5305)
w12	0.9	0.8918	0.0451	(0.8048 , 0.9822)	0.8509	0.0426	(0.7683 , 0.9351)
w15	0.5	0.5013	0.0736	(0.3522 , 0.6414)			
w22	2	2.0128	0.0449	(1.9274 , 2.1031)	1.9499	0.0418	(1.8695 , 2.0339)
w25	0.7	0.7418	0.0459	(0.6502 , 0.8302)			
w33	1.8	1.8222	0.0427	(1.7405 , 1.9083)	1.8001	0.0417	(1.7204 , 1.8836)
w34	0.6	0.6135	0.0322	(0.5516 , 0.6774)	0.5793	0.0311	(0.5189 , 0.6406)
w35	0.6	0.6249	0.0501	(0.5263 , 0.7220)			
w44	2.5	2.4472	0.0471	(2.3570 , 2.5413)	2.4081	0.0456	(2.3207 , 2.4994)
w45	0.7	0.6805	0.0501	(0.5834 , 0.7777)			

Figure 1

M_0: Posterior Distributions for selected parameters



(a). Posterior distribution for $\beta_{10}(1)$



(b). Posterior Distribution for ω_{13}

Table 2

M_2 and M_3: Estimation for generated data

	M_2 Unrestricted			M_3 Restricted		
	Mean	SD	95% Credible Interval	Mean	SD	95% Credible Interval
Const1	-1.0201	0.0487	(-1.1168 , -0.9252)	-1.3382	0.0464	(-1.4309, -1.2486)
ν_1	0.1818	0.0152	(0.1521, 0.2115)	0.1807	0.0152	(0.1511, 0.2107)
d	-1.0543	0.0783	(-1.2085 , -0.9024)	-0.2907	0.0643	(-0.4165 , -0.1660)
Const2	1.0324	0.0558	(0.9240 , 1.1423)	0.7229	0.0509	(0.6236, 0.8232)
ν_1	-0.1749	0.0192	(-0.2122 , -0.1370)	-0.1754	0.0192	(-0.2132, -0.1376)
ν_2	-0.2507	0.0184	(-0.2866, -0.2150)	-0.2431	0.0186	(-0.2793, -0.2063)
d	-1.9571	0.0959	(-2.1464 , -1.7689)	-1.2208	0.0781	(-1.3739, -1.0674)
Const3	-0.4539	0.0194	(-0.4916, -0.4158)	-0.4681	0.0202	(-0.5073, -0.4286)
ν_1	1.0088	0.0193	(0.9710, 1.0467)	1.0068	0.0194	(0.9690, 1.0455)
ν_2	-0.4935	0.0119	(-0.5171, -0.4709)	-0.5025	0.0126	(-0.5277, -0.4779)
ω_{11}	5.8743	0.1715	(5.5540 , 6.2277)	5.7140	0.1531	(5.4203, 6.0192)
ω_{12}	-2.4337	0.1272	(-2.6867, -2.1886)	-2.5701	0.1235	(-2.8140, -2.3263)
ω_{13}	1.1753	0.0640	(1.0497 , 1.2994)			
ω_{22}	15.0195	0.2151	(14.5992 , 15.4459)	14.8819	0.2110	(14.4721, 15.3030)
ω_{23}	1.1307	0.0815	(0.9702 , 1.2897)			

Table 3

Log Marginal Likelihood

			Gauss-Lagrange		Importance Sampling	
Model	prior	posterior	likelihood	marginal likelihood	likelihood	marginal likelihood
M_0	-46.90	56.74	-52071.25	-52174.89	-52071.37	-52175.01
M_1	-44.41	47.90	-52275.76	-52368.03	-52276.24	-52368.51
M_2	-37.32	35.76	-59400.55	-59473.64	-59400.10	-59473.18
M_3	-34.92	31.75	-59505.26	-59571.93	-59505.18	-59571.85

Please note that all numbers are in logarithms.

5 Empirical Application

We applied our model to the similar applications as Munkin and Trivedi (2003) did. The first data sample is obtained from the 1987-1988 National Medical Expenditure Survey(NMES, 1987), which contains U.S. elderly population data with positive medical expenditures. The second sample is obtained from the 1996 Medical Expenditure Panel Survey (MEPS), which consists of non-elderly privately insured individuals with positive medical expenditures. The sample used in our paper is slightly different from the sample used in Munkin and Trivedi (2003), because the sample used in their paper is not publicly available.

The summary of statistics and variable definitions are presented in Table 4.

5.1 Private Insurance

The first application is to study the impact of private insurance on the number of physician doctor office-based visits and the associated office-based total expenditures. We obtained 3680 elderly Americans observations from the National Medical Expenditure Survey (NMES, 1987), and the summary statistics of our sample is presented in Table 4.

Individuals older than 65 are covered by Medicare, which covers the costs of treatment of a wide range of health care services. Some individuals choose to purchase the private insurance if their health conditions are bad, especially if they have chronic conditions. We assume that by controlling all the other covariates, including family income, self-perceived health conditions, number of chronic conditions, education level, employment, etc., the private insurance is exogenous.

We use our model to analyze the impact of private insurance on the number of physician office-based visits and associated expenditures. Following Munkin and Trivedi (2003), we use self-perceived health status variables, number of chronic conditions, disability status, location, demographic variables, and insurance variables as the covariates in equations that determine the number of physician office visits and health care expenditures. Even though that there may be heterogeneous effects due to different types of the private insurance policies, the main focus here is to study the impact of the Medigap plans. Family income is assumed to influence the purchase of the private insurance, while it would not affect the health care utilization directly. Medicaid is assumed to have no impact on the selection equation.

The posterior mean and standard deviation are summarized in Table 5 and Table 6, and the 95% credible interval statistics are summarized in Table 7. The prior settings are the same as in the previous section. We estimated the parameters using 50,000 iterations. It is noticeable that the standard deviation for the control group is larger than what we have obtained for the treatment group. This is due to the fact that only 20% of the sample is in the control group. Because of the larger standard deviations in the control group, there are many coefficient which are not substantially different from zero in the control group, are

Table 4

Variable Definition and Summary Statistics

Variable	Data set	NMES		MEPS	
	Number of Observations	3680		5368	
	Definition	Mean	St.Dev	Mean	St.Dev
DOCVIS	# of physician office visits	6.88	6.87	4.88	5.98
DVEXP	Expenditure on physician office visits	422.8	785.5	499.5	1047.5
EXCHLTH	=1 if self-perceived health is excellent	0.09	0.29	0.32	0.47
POORHLTH	=1 if self-perceived health is poor	0.14	0.34	0.02	0.15
NUMCHRON	# of chronic conditions	2.02	1.42	0.80	1.12
ADLDIFF	=1 if has a condition which limits activities of daily living	0.21	0.40		
INJURY	# of injuries which limit activities of daily living since 1996			0.41	0.82
NOREAST	=1 if lives in northeastern U.S.	0.19	0.39	0.21	0.40
MIDWEST	=1 if lives in midwestern U.S.	0.26	0.44	0.25	0.43
WEST	=1 if lives in western U.S.	0.19	0.39	0.21	0.41
AGE	age in years (divided by 10)	7.40	0.62	4.14	1.25
BLACK	=1 if is African American	0.10	0.31	0.10	0.30
FEMALE	= 1 if female	0.61	0.49	0.58	0.49
MARRIED	= 1 if the person is married	0.56	0.50	0.68	0.47
SCHOOL	# of years of education	10.6	3.5	13.32	2.58
FAMINC	Family income in $1,000	25.8	30.1	59.11	39.02
EMPLOYED	=1 if the person is employed	0.10	0.30	0.82	0.38
PRIVATE	=1 if covered by private health insurance	0.80	0.40		
MEDICAID	=1 if covered by Medicaid	0.09	0.28		
INSURANCE	=0 if covered by HMO = 1 if FFS			0.51	0.50
SELFEMP	=1 if self-employed			0.09	0.28
SIZE	The size of the company where the person works			127.5	177.8
LOCATION	=1 if the company has multiple locations			0.52	0.50
GOVT	= 1 if the company is governmental			0.18	0.39

significant in the treated group, including EXCHLTH, NOREAST and SCHOOL in both the doctor's visit and expenditure equations. The estimates for the coefficients of some other variables, inclduing ADLDIFF, and MEDICAID are significant in the control group but not in the treated group.

Based on the 95% credible interval results, there is no substantial difference between the posterior means for majority of the coefficients in control and treatment group. The only exception are EXCHLTH, BLACK, ADLDIFF and SCHOOL in the expenditure equation, and SCHOOL in the doctor's visit equation. The 95% credible interval for the coefficient in the treatment group does not overlap with that of the control group. We estimated another model, denoted by M_4, by restricting all the other covariates having the same marginal impact for both groups, and allowing different marginal impact of these variables with flexible intercepts. The estimated results are summarized in Table 8. Based on the results, we found that there are no significant differences between the coefficients for those variables anymore.

Based on the results in M_1 and M_4, we can obtain some evidence indicating the existence of selection bias. By taking in to consideration of the covariance between the errors in the health care utilization equations and the selection equation, the impact of private insurance purchase is no longer significant.

The result comparison between our parsimonious models M_2 and M_3 with the results from Munkin and Trivedi (2003) is presented in Table 6. The coefficients estimated for PRIVINS in model M_3 are larger than those of M_2, and the standard deviations are smaller. By taking into the consideration of the correlation between the healthcare utilization with the selection decision, the impact of private insurance purchase decision is no longer significant. However, the covariance estimates of $\omega_{1\mu}$ and $\omega_{2\mu}$ is not significantly different from zero. Thus only weak evidence of selection bias is detected.

The marginal log-likelihood calculated using the importance sampling is summarized in Table 13. Note that we assume that $p(M_i) = p(M_j) = \frac{1}{2}$ when we are calculating the Bayes factor for comparing model M_i and M_j. The winning model is M_3, the restricted constant treatment model. This suggest that when dealing with this sample, the constant treatment model is recommended.

Table 5

M_0 vs. M_1: MCMC estimates of the private insurance model

	M0					M1				
	INS	VIS(C)	EXP(C)	VIS	EXP	INS	VIS(C)	EXP(C)	VIS	EXP
CONST	-0.1812	1.6879	4.6143	1.0655	4.2090	-0.2317	1.8203	5.0835	1.2324	4.6382
	0.3390	0.4111	0.5793	0.2411	0.3304	0.3539	0.3951	0.5284	0.2265	0.2979
EXCHLTH	0.0693	0.0187	-0.0159	-0.1941	-0.1842	0.0602	0.0309	0.0192	-0.1957	-0.1883
	0.0999	0.1444	0.1958	0.0558	0.0732	0.1036	0.1426	0.1884	0.0550	0.0703
POORHLTH	-0.0494	0.3205	0.3566	0.3241	0.3770	-0.0600	0.3180	0.3310	0.3346	0.4062
	0.0712	0.0822	0.1181	0.0499	0.0690	0.0738	0.0803	0.1106	0.0487	0.0668
NUMCHRON	-0.0334	0.1375	0.1199	0.1350	0.1298	-0.0367	0.1317	0.1012	0.1380	0.1384
	0.0185	0.0227	0.0324	0.0121	0.0164	0.0188	0.0223	0.0299	0.0118	0.0158
ADLDIFF	-0.2087	0.1686	0.3456	0.0620	-0.0021	-0.2424	0.1383	0.2340	0.0866	0.0654
	0.0659	0.0787	0.1138	0.0462	0.0645	0.0661	0.0747	0.1014	0.0442	0.0593
NOREAST	0.1241	0.0646	0.0821	0.0997	0.2655	0.1434	0.0884	0.1565	0.0857	0.2258
	0.0732	0.0953	0.1331	0.0457	0.0618	0.0746	0.0932	0.1239	0.0441	0.0587
MIDWEST	0.2765	-0.0838	-0.2666	0.0322	0.0193	0.2876	-0.0372	-0.0990	0.0062	-0.0440
	0.0681	0.1006	0.1409	0.0420	0.0577	0.0714	0.0955	0.1268	0.0399	0.0526
WEST	-0.1447	0.2202	0.5217	0.0544	0.2719	-0.1697	0.1959	0.4278	0.0660	0.2969
	0.0724	0.0946	0.1349	0.0462	0.0626	0.0752	0.0918	0.1246	0.0449	0.0600
BLACK	-0.7995	0.0997	0.3893	-0.2247	-0.4146	-0.8475	-0.0270	-0.0558	-0.1143	-0.1369
	0.0756	0.1035	0.1734	0.0873	0.1213	0.0757	0.0732	0.0985	0.0693	0.0911
MALE	-0.0086	-0.0094	-0.0741	-0.0209	0.0106	-0.0368	-0.0157	-0.0920	-0.0160	0.0212
	0.0585	0.0814	0.1101	0.0356	0.0480	0.0613	0.0799	0.1054	0.0349	0.0462
MARRIED	0.2524	-0.0771	-0.1725	-0.0047	0.0613	0.2645	-0.0288	0.0005	-0.0332	-0.0071
	0.0587	0.0843	0.1208	0.0380	0.0528	0.0614	0.0782	0.1042	0.0357	0.0472
SCHOOL	0.1057	-0.0120	-0.0391	0.0299	0.0540	0.1079	0.0060	0.0247	0.0206	0.0296
	0.0082	0.0144	0.0241	0.0068	0.0099	0.0084	0.0098	0.0132	0.0051	0.0067
AGE	-0.0088	-0.1098	-0.0912	-0.0179	-0.0002	-0.0048	-0.1088	-0.0843	-0.0168	0.0033
	0.0426	0.0514	0.0713	0.0282	0.0380	0.0444	0.0502	0.0672	0.0280	0.0371
EMPLOYED	0.0912	0.0749	-0.0623	-0.0160	-0.0069	0.0832	0.1023	0.0391	-0.0238	-0.0226
	0.0912	0.1440	0.1968	0.0521	0.0702	0.0991	0.1405	0.1869	0.0512	0.0674
MEDICAID		0.1673	0.2142	0.1948	0.2054		0.1627	0.1997	0.2026	0.2274
		0.0752	0.1003	0.1191	0.1583		0.0741	0.1000	0.1210	0.1644
FAMINC	0.0043					0.0061				
	0.0013					0.0014				
ω_{13}, ω_{23}		-0.2172	-0.8033	0.2316	0.6184					
		0.1294	0.2429	0.1362	0.1734					
ω_{11}, ω_{22}		0.5933	1.8003	0.5548	1.3751		0.5361	1.3142	0.5252	1.2422
		0.0618	0.2634	0.0268	0.0696		0.0403	0.0699	0.0184	0.0324
ω_{12}			0.8716		0.7582			0.7278		0.7116
			0.1148		0.0292			0.0469		0.0219

Table 6

M2 vs. M3: MCMC estimates of the private insurance model

	M2			M3		
	INS	DOCVIS(C)	DVEXP(C)	INS	DOCVIS(C)	DVEXP(C)
CONST	-0.2350	1.3162	4.5915	-0.2326	1.2678	4.5512
	0.3544	0.2352	0.2950	0.3520	0.1986	0.2632
EXCHLTH	0.0570	-0.1674	-0.1656	0.0589	-0.1685	-0.1656
	0.1027	0.0517	0.0663	0.1026	0.0513	0.0657
POORHLTH	-0.0584	0.3233	0.3788	-0.0582	0.3251	0.3807
	0.0742	0.0420	0.0572	0.0736	0.0412	0.0568
NUMCHRON	-0.0361	0.1351	0.1283	-0.0368	0.1356	0.1288
	0.0192	0.0106	0.0142	0.0190	0.0103	0.0139
ADLDIFF	-0.2387	0.0918	0.1042	-0.2419	0.0983	0.1094
	0.0665	0.0411	0.0540	0.0664	0.0379	0.0513
MEDICAID		0.1795	0.2257		0.1862	0.2321
		0.0602	0.0804		0.0591	0.0789
PRIVINS		**0.0868**	**0.2311**		**0.1838**	**0.3157**
		0.2372	**0.2621**		**0.0434**	**0.0573**
NOREAST	0.1422	0.0870	0.2175	0.1432	0.0819	0.2135
	0.0747	0.0414	0.0545	0.0751	0.0396	0.0527
MIDWEST	0.2876	0.0087	-0.0428	0.2877	0.0026	-0.0489
	0.0706	0.0400	0.0516	0.0715	0.0368	0.0485
WEST	-0.1695	0.0910	0.3233	-0.1691	0.0927	0.3252
	0.0754	0.0413	0.0546	0.0755	0.0404	0.0541
BLACK	-0.8441	-0.0985	-0.1251	-0.8474	-0.0716	-0.1010
	0.0766	0.0830	0.0988	0.0766	0.0495	0.0654
MALE	-0.0348	-0.0138	0.0035	-0.0361	-0.0133	0.0039
	0.0606	0.0319	0.0424	0.0614	0.0320	0.0424
MARRIED	0.2627	-0.0281	-0.0035	0.2647	-0.0342	-0.0091
	0.0620	0.0364	0.0469	0.0614	0.0325	0.0432
SCHOOL	0.1079	0.0196	0.0304	0.1079	0.0172	0.0281
	0.0084	0.0077	0.0091	0.0084	0.0045	0.0060
AGE	-0.0046	-0.0413	-0.0227	-0.0048	-0.0420	-0.0232
	0.0444	0.0248	0.0328	0.0441	0.0246	0.0326
EMPLOYED	0.0833	0.0749	-0.0233	0.0831	-0.0185	-0.0252
	0.0978	0.0483	0.0636	0.0989	0.0482	0.0636
FAMINC	0.0060			0.0061		
	0.0014			0.0015		
ω_{13}, ω_{23}		0.0535	0.0474			
		0.1302	0.1432			
ω_{11}, ω_{22}		0.5342	1.2650		0.5265	1.2556
		0.0196	0.0319		0.0166	0.0294
ω_{12}			0.7220			0.7145
			0.0224			0.0198

Table 7

95% Credible Interval for the private insurance model M_0

	INSURANCE		DOCVIS(C)		DVEXP(C)		DOCVIS		DVEXP	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
CONST	-0.8481	0.4781	0.8788	2.4927	3.4713	5.7404	0.5846	1.5288	3.5647	4.8625
EXCHLTH	-0.1227	0.2676	-0.2659	0.3024	**-0.4005**	**0.3668**	-0.3035	-0.0849	-0.3281	-0.0407
POORHLTH	-0.1898	0.0909	0.1602	0.4806	0.1251	0.5881	0.2264	0.4217	0.2426	0.5126
NUMCHRON	-0.0695	0.0030	0.0932	0.1825	0.0569	0.1836	0.1113	0.1585	0.0978	0.1618
ADLDIFF	-0.3380	-0.0806	0.0150	0.3218	**0.1247**	**0.5702**	-0.0286	0.1520	-0.1277	0.1246
NOREAST	-0.0187	0.2672	-0.1217	0.2506	-0.1791	0.3433	0.0104	0.1899	0.1455	0.3869
MIDWEST	0.1437	0.4098	-0.2812	0.1118	-0.5441	0.0085	-0.0494	0.1149	-0.0928	0.1324
WEST	-0.2868	-0.0037	0.0350	0.4044	0.2571	0.7869	-0.0361	0.1455	0.1485	0.3949
BLACK	-0.9481	-0.6514	-0.1043	0.3011	**0.0131**	**0.7032**	-0.4046	-0.0616	-0.6345	-0.1592
MALE	-0.1235	0.1058	-0.1689	0.1499	-0.2911	0.1409	-0.0913	0.0491	-0.0827	0.1048
MARRIED	0.1372	0.3673	-0.2429	0.0873	-0.4080	0.0662	-0.0792	0.0702	-0.0426	0.1640
SCHOOL	0.0899	0.1219	**-0.0397**	**0.0172**	**-0.0821**	**0.0144**	0.0174	0.0440	0.0330	0.0714
AGE	-0.0923	0.0747	-0.2110	-0.0096	-0.2309	0.0488	-0.0730	0.0377	-0.0747	0.0743
EMPLOYED	-0.0969	0.2810	-0.2068	0.3560	-0.4514	0.3239	-0.1181	0.0861	-0.1444	0.1311
MEDICAID			0.0204	0.3145	0.0178	0.4119	-0.0386	0.4288	-0.1038	0.5164
FAMINC	0.0019	0.0070								

The bold text indicates that the 95% credible interval doesn't overlap with each other for the control and treatment group.

5.2 HMO and FFS

We also compare the impact of choosing different types of private insurances on the health care utilization. We generated a comparable sample of 5368 observations. The summary of statistics is presented in Table 4.

Similar to the previous section, we assume that the employment variables, including company size, location, self-employment indicator and family income only affect the private insurance selection, and has no direct impact on the health-care utilization. This assumption is realistic in the sense that the majority of the individuals have limited choice to the health insurance plans, and those choices are closely related to their job.

The posterior mean and standard deviation are summarized in Table 9 and Table 10, and the 95% credible interval statistics for M_0 are summarized in Table 11. The estimates for the coefficients of EXCHLTH, WEST, and AGE in the treatment group are significantly different from those of the control group in the expenditure equation. We estimated a another model M_4 by assuming that only these covariates can interact with the treatment variable and

Table 8

M4: MCMC estimates of the private insurance model

	M4				
	INS	VIS(C)	EXP(C)	VIS	EXP
CONST	-0.2436	1.4172	4.6748	1.2607	4.3119
	0.3382	0.2194	0.3116	0.2047	0.2811
EXCHLTH	0.0645	-0.1655	-0.1369		-0.1660
	0.0985	0.0519	0.1402		0.0716
POORHLTH	-0.0526	0.3216	0.3570		
	0.0708	0.0423	0.0594		
NUMCHRON	-0.0350	0.1336	0.1206		
	0.0184	0.0105	0.0146		
ADLDIFF	-0.1975	0.0759	0.1609		-0.0039
	0.0643	0.0390	0.0787		0.0574
NOREAST	0.1072	0.0970	0.2564		
	0.0720	0.0401	0.0560		
MIDWEST	0.2689	0.0266	0.0209		
	0.0678	0.0376	0.0517		
WEST	-0.1606	0.0830	0.2933		
	0.0731	0.0411	0.0571		
BLACK	-0.8131	-0.1655	-0.2499		-0.4332
	0.0746	0.0641	0.1262		0.0913
MALE	-0.0220	-0.0186	-0.0085		
	0.0581	0.0325	0.0442		
MARRIED	0.2496	-0.0086	0.0683		
	0.0592	0.0338	0.0463		
SCHOOL	0.1068	0.0255	0.0523	0.0261	0.0563
	0.0081	0.0147	0.0214	0.0055	0.0074
AGE	-0.0003	-0.0371	-0.0179		
	0.0423	0.0247	0.0341		
EMPLOYED	0.0948	-0.0087	-0.0034		
	0.0946	0.0487	0.0666		
MEDICAID		0.1674	0.2078		
		0.0615	0.0819		
FAMINC	0.0044				
	0.0013				
ω_{13}, ω_{23}		0.1913	0.3024	0.1546	0.7006
		0.1291	0.2057	0.0463	0.0657
ω_{11}, ω_{22}		0.5727	1.4062	0.5452	1.3995
		0.0558	0.1152	0.0204	0.0489
ω_{12}			0.7813		0.7550
			0.0730		0.0288

all the other covariates has same marginal impact in both the treated and contorl group. The results are presented in Table 12. Based on the results, only the variable WEST has significantly different impact on the expenditure equation for the control group and treatment group. The covariance coefficients in the treated group are significantly different from zero. Note that based on the results in M_1, only the coefficients for the variable INJURY are significantly different in the expenditure equation. This suggested some evidence of selection bias, as the impact of treatment are different if we considering the existence of the correlation between healthcare utilization and the selection decision.

Based on the credible interval results, the estimated covariance parameters between the errors of the doctor's visit equation and the selection equation is negative in the control group, but positive in the treated group. Similar patterns were identified for the estimated covariance parameters between the errors of the expenditure equation and the selection equation.

The comparison between the estimates from model M_2 and M_3 is presented in Table 10. The standard deviation for the coefficient estimates of the variable INSURANCE are smaller in the restricted model. The impact of INSURANCE in both M_2 and M_3 is not significantly different from zero.

According to the results in Table 13, the selected model is M_4, which suggested that there is some evidence supporting the existence of endogeneity of the treatment model.

5.3 Conclusion

To conclude, we proposed a parametric self-selection model with multiple outcomes in Bayesian settings to estimate the impact of health insurance on health care expenditures. There are two outcomes in our model: one is the number of doctor's office visits, and the other is the health care expenditure. We applied the model to two empirical applications, and compared the results with the results using the other four more parsimonious models. We found some evidence supporting the selection bias in both applications.

Table 9

M_0 and M_1: MCMC estimates of the HMO model

	M0 Unrestricted					M1 Restricted				
	INS	VIS(C)	EXP(C)	VIS	EXP	INS	VIS(C)	EXP(C)	VIS	EXP
CONST	-0.0211	0.2362	3.4330	0.2749	3.3497	0.0043	0.4679	4.2982	0.4473	4.1346
	0.1142	0.1333	0.1788	0.1268	0.1690	0.1189	0.1203	0.1536	0.1165	0.1467
EXCLHLTH	0.0863	-0.2479	-0.2217	-0.1658	-0.0047	0.1103	-0.2324	-0.1445	-0.1826	-0.0692
	0.0375	0.0411	0.0562	0.0400	0.0536	0.0387	0.0401	0.0507	0.0393	0.0488
POORHLTH	0.0033	0.4081	0.4297	0.2278	0.3388	0.0057	0.4204	0.4733	0.2397	0.3506
	0.1156	0.1169	0.1788	0.1060	0.1586	0.1187	0.1139	0.1629	0.1038	0.1437
NUMCHRON	0.0135	0.2169	0.2190	0.2311	0.2734	0.0060	0.2187	0.2229	0.2294	0.2682
	0.0162	0.0170	0.0252	0.0150	0.0223	0.0169	0.0165	0.0230	0.0146	0.0202
INJURY	-0.0009	0.1749	0.2294	0.1193	0.1101	0.0138	0.1776	0.2460	0.1193	0.1063
	0.0211	0.0228	0.0335	0.0194	0.0290	0.0220	0.0222	0.0307	0.0191	0.0259
NOREAST	-0.0867	0.1057	0.0644	0.0970	0.0840	-0.1168	0.0864	-0.0276	0.1124	0.1511
	0.0476	0.0499	0.0705	0.0487	0.0683	0.0489	0.0482	0.0632	0.0478	0.0621
MIDWEST	0.1306	-0.0625	-0.2809	-0.0285	-0.0559	0.1546	-0.0409	-0.1956	-0.0511	-0.1576
	0.0452	0.0511	0.0703	0.0451	0.0624	0.0464	0.0497	0.0645	0.0432	0.0557
WEST	-0.3861	-0.0018	0.2398	-0.0876	-0.2798	-0.4259	-0.0718	-0.0275	-0.0336	-0.0299
	0.0481	0.0508	0.0702	0.0557	0.0755	0.0495	0.0471	0.0609	0.0524	0.0678
BLACK	-0.2259	-0.1568	-0.1160	-0.1623	-0.3319	-0.2143	-0.1994	-0.2785	-0.1244	-0.1717
	0.0593	0.0598	0.0831	0.0678	0.0904	0.0609	0.0578	0.0736	0.0657	0.0831
FEMALE	-0.0556	0.3711	0.3999	0.3026	0.3335	-0.0703	0.3625	0.3548	0.3079	0.3694
	0.0348	0.0373	0.0519	0.0362	0.0496	0.0361	0.0366	0.0469	0.0356	0.0452
MARRIED	0.0385	-0.0676	-0.1100	0.0065	0.0116	0.0229	-0.0610	-0.0949	0.0032	0.0021
	0.0390	0.0392	0.0558	0.0405	0.0556	0.0404	0.0384	0.0504	0.0398	0.0509
SCHOOL	0.0039	0.0305	0.0383	0.0245	0.0368	0.0028	0.0298	0.0358	0.0243	0.0383
	0.0069	0.0071	0.0101	0.0071	0.0096	0.0072	0.0069	0.0091	0.0069	0.0088
EMPLOYED	-0.1768	-0.0385	0.0905	-0.1382	-0.1801	-0.1440	-0.0829	-0.0759	-0.1021	-0.0263
	0.0528	0.0529	0.0741	0.0454	0.0647	0.0585	0.0508	0.0671	0.0442	0.0575
AGE	0.0682	0.0099	0.0362	0.0505	0.1337	0.0786	0.0211	0.0847	0.0398	0.0829
	0.0151	0.0166	0.0229	0.0154	0.0215	0.0157	0.0158	0.0206	0.0150	0.0194
FAMINC	-0.0008					-0.0009				
	0.0004					0.0005				
SELFEMP	0.0775					0.1222				
	0.0539					0.0733				
GOVT	0.0362					-0.0205				
	0.0358					0.0483				
SIZE	-0.0004					-0.0006				
	0.0001					0.0001				
LOCATION	-0.0193					-0.0525				
	0.0339					0.0462				
$\omega_{1\mu}, \omega_{2\mu}$		-0.2669	-1.0445	0.1977	0.9768					
		0.0554	0.0651	0.0542	0.0677					
ω_{11}, ω_{22}		0.5987	2.0157	0.5910	1.9004		0.5322	1.3450	0.5449	1.3231
		0.0312	0.1040	0.0277	0.0988		0.0212	0.0371	0.0212	0.0358
ω_{12}			0.9353		0.8773			0.7432		0.7403
			0.0590		0.0533			0.0249		0.0245

Table 10

M2 vs. M3: MCMC estiamtes of the HMO model

	M2 Unrestricted			M3 Restricted		
	INS	DOCVIS(C)	DVEXP(C)	INS	DOCVIS(C)	DVEXP(C)
CONST	0.0015	0.5298	4.2901	0.0040	0.4625	4.1925
	0.1193	0.1496	0.1867	0.1188	0.0840	0.1075
EXCLHLTH	0.1096	-0.2024	-0.0964	0.1107	-0.2079	-0.1045
	0.0387	0.0300	0.0378	0.0387	0.0283	0.0351
POORHLTH	0.0063	0.3301	0.4077	0.0051	0.3281	0.4058
	0.1197	0.0774	0.1094	0.1195	0.0766	0.1082
NUMCHRON	0.0062	0.2247	0.2488	0.0062	0.2242	0.2482
	0.0168	0.0112	0.0155	0.0169	0.0110	0.0152
INJURY	0.0138	0.1449	0.1655	0.0140	0.1440	0.1641
	0.0219	0.0147	0.0201	0.0221	0.0145	0.0199
INSURANCE		**-0.1445**	**-0.1581**		**-0.0044**	**0.0347**
		0.2442	**0.2989**		**0.0248**	**0.0321**
NOREAST	-0.1168	0.0935	0.0535	-0.1169	0.1007	0.0629
	0.0485	0.0369	0.0471	0.0484	0.0342	0.0444
MIDWEST	0.1526	-0.0390	-0.1670	0.1540	-0.0482	-0.1784
	0.0464	0.0358	0.0457	0.0462	0.0326	0.0421
WEST	-0.4250	-0.0780	-0.0509	-0.4262	-0.0555	-0.0204
	0.0493	0.0521	0.0659	0.0495	0.0349	0.0447
BLACK	-0.2147	-0.1750	-0.2510	-0.2145	-0.1624	-0.2320
	0.0612	0.0490	0.0626	0.0608	0.0428	0.0549
FEMALE	-0.0688	0.3309	0.3583	-0.0699	0.3336	0.3631
	0.0362	0.0266	0.0337	0.0359	0.0254	0.0325
MARRIED	0.0233	-0.0301	-0.0487	0.0229	-0.0306	-0.0506
	0.0402	0.0281	0.0363	0.0405	0.0275	0.0356
SCHOOL	0.0030	0.0269	0.0367	0.0028	0.0270	0.0369
	0.0072	0.0050	0.0064	0.0072	0.0049	0.0064
EMPLOYED	-0.1442	-0.1084	-0.0659	-0.1444	-0.0953	-0.0467
	0.0584	0.0411	0.0532	0.0586	0.0332	0.0434
AGE	0.0789	0.0366	0.0918	0.0787	0.0320	0.0859
	0.0156	0.0134	0.0170	0.0158	0.0109	0.0141
FAMINC	-0.0009			-0.0009		
	0.0005			0.0005		
SELFEMP	0.1211			0.1227		
	0.0722			0.0721		
GOVT	-0.0228			-0.0212		
	0.0487			0.0482		
SIZE	-0.0006			-0.0006		
	0.0001			0.0001		
LOCATION	-0.0600			-0.0525		
	0.0470			0.0460		
$\omega_{1\mu}, \omega_{2\mu}$		0.0871	0.1208			
		0.1518	0.1862			
ω_{11}, ω_{22}		0.5572	1.3674		0.5381	1.3370
		0.0299	0.0463		0.0151	0.0257
ω_{12}			0.7618			0.7418
			0.0330			0.0175

Table 11

95% Credible Interval for the HMO model M_0

	INSURANCE		DOCVIS(C)		DVEXP(C)		DOCVIS		DVEXP	
	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper	Lower	Upper
CONST	-0.2459	0.2040	-0.0258	0.4950	3.0789	3.7797	0.0258	0.5235	3.0131	3.6765
EXCLHLTH	0.0121	0.1601	-0.3291	-0.1679	**-0.3311**	**-0.1110**	-0.2445	-0.0879	-0.1094	0.1008
POORHLTH	-0.2237	0.2300	0.1785	0.6379	0.0784	0.7797	0.0205	0.4357	0.0307	0.6478
NUMCHRON	-0.0180	0.0451	0.1838	0.2501	0.1697	0.2687	0.2017	0.2605	0.2294	0.3171
INJURY	-0.0422	0.0408	0.1302	0.2199	0.1640	0.2950	0.0812	0.1572	0.0532	0.1666
NOREAST	-0.1798	0.0061	0.0083	0.2035	-0.0725	0.2020	0.0016	0.1917	-0.0492	0.2183
MIDWEST	0.0420	0.2191	-0.1640	0.0372	-0.4196	-0.1430	-0.1171	0.0596	-0.1772	0.0670
WEST	-0.4805	-0.2920	-0.1018	0.0974	**0.1027**	**0.3783**	-0.1976	0.0206	-0.4292	-0.1327
BLACK	-0.3430	-0.1098	-0.2733	-0.0394	-0.2790	0.0481	-0.2949	-0.0293	-0.5087	-0.1539
FEMALE	-0.1238	0.0124	0.2982	0.4448	0.2984	0.5024	0.2317	0.3732	0.2356	0.4308
MARRIED	-0.0383	0.1153	-0.1442	0.0088	-0.2189	-0.0002	-0.0737	0.0855	-0.0971	0.1205
SCHOOL	-0.0096	0.0174	0.0166	0.0446	0.0184	0.0579	0.0105	0.0382	0.0179	0.0555
EMPLOYED	-0.2802	-0.0740	-0.1409	0.0657	-0.0539	0.2372	-0.2263	-0.0487	-0.3064	-0.0536
AGE	0.0387	0.0979	-0.0221	0.0428	**-0.0089**	**0.0809**	0.0202	0.0808	0.0918	0.1762
FAMINC	-0.0015	-0.0001								
SELFEMP	-0.0286	0.1831								
GOVT	-0.0332	0.1066								
SIZE	-0.0005	-0.0002								
LOCATION	-0.0850	0.0475								

The bold text indicates that the 95% credible interval doesn't overlap with each other for the control and treatment group.

Table 12

M4: MCMC estiamtes of the HMO model

	M4				
	INS	VIS(C)	EXP(C)	VIS	EXP
CONST	0.0253	-0.0070	3.8995	0.3586	3.5001
	0.1076	0.1635	0.1931	0.0958	0.1289
EXCLHLTH	0.0841	-0.2161	-0.1209		-0.0519
	0.0364	0.0287	0.0440		0.0440
POORHLTH	-0.0058	0.3053	0.3935		
	0.1073	0.0779	0.1094		
NUMCHRON	-0.0064	0.2234	0.2461		
	0.0150	0.0112	0.0153		
INJURY	0.0307	0.1374	0.1704		
	0.0192	0.0148	0.0201		
NOREAST	-0.1253	0.1162	0.0413		
	0.0431	0.0358	0.0452		
MIDWEST	0.0902	-0.0551	-0.1445		
	0.0418	0.0338	0.0432		
WEST	-0.3932	0.0048	0.0590		-0.2203
	0.0472	0.0436	0.0613		0.0643
BLACK	-0.1937	-0.1240	-0.2482		
	0.0541	0.0470	0.0583		
FEMALE	-0.0632	0.3417	0.3500		
	0.0321	0.0260	0.0331		
MARRIED	0.0213	-0.0303	-0.0526		
	0.0356	0.0283	0.0358		
SCHOOL	0.0025	0.0271	0.0357		
	0.0064	0.0050	0.0063		
EMPLOYED	-0.1258	-0.0761	-0.0946		
	0.0511	0.0357	0.0468		
AGE	0.0683	0.0253	0.0848		0.1162
	0.0146	0.0116	0.0181		0.0175
FAMINC	-0.0011				
	0.0004				
SELFEMP	0.0563				
	0.0588				
GOVT	0.0317				
	0.0400				
SIZE	-0.0004				
	0.0001				
LOCATION	-0.0370				
	0.0383				
$\omega_{1\mu}, \omega_{2\mu}$		-0.5961	-0.4427	0.1189	0.8870
		0.1754	0.1961	0.0567	0.0722
ω_{11}, ω_{22}		0.7410	1.4886	0.5758	1.7906
		0.0753	0.0842	0.0241	0.0942
ω_{12}			0.9079		0.8182
			0.0774		0.0482

Table 13

Log Marginal Likelihood

	NMES		MEPS	
Model	Gauss-Lagrange	Importance Sampling	Gauss-Lagrange	Importance Sampling
M_0	-16212.75	-16212.64	-23511.57	-23511.46
M_1	-16236.18	-16236.19	-23589.14	-23588.96
M_2	-16133.89	-16134.04	-23487.98	-23487.90
M_3	**-16130.36**	**-16130.49**	-23485.37	-23485.48
M_4	-16145.97	-16145.86	**-23463.98**	**-23462.49**

Appendix A M_2: Model Specification

We observe N independent observations and use y_{i1} and y_{i2} to denote the outcome variables, where y_{i1} is the discrete potential outcome variable and y_{i2} is the potential continuous outcome variable, $k \in \{0, 1\}$ to denote the treatment status. Let x_{i1} denote the exogenous covariates matrix for y_1 and x_{i2} denote the exogenous covariates matrix for y_2 for individual i. We assume that

$$y_{i1} \sim \text{Poisson}(\mu_i) \tag{17}$$

Let's use T_i to denote the treatment variable for subject i.

z_i is a latent variable that relates to T_i, and $T_i = \mathbb{1}\{z_i \geq 0\}$.

The model can be represented represented as following equation:

$$y_i = X_i\beta + \epsilon_i \tag{18}$$

where

$$y_i = (\ln(\mu_i), y_{i2}, z_i)', \quad X_i = \begin{pmatrix} x'_{i1} & 0 & 0 \\ 0 & x'_{i2} & 0 \\ 0 & 0 & x'_{i3} \end{pmatrix}$$

$\beta = (\beta'_1, \beta'_2, \beta'_3)'$ and $\epsilon_i = (\epsilon_{i1}, \epsilon_{i2}, \epsilon_{i3})'$.

Assume that $\epsilon_i \sim N(0, \Omega)$. Let's define:

$$\Omega = \begin{pmatrix} \omega_{11} & \omega_{12} & \omega_{13} \\ \omega_{21} & \omega_{22} & \omega_{23} \\ \omega_{31} & \omega_{32} & 1 \end{pmatrix}$$

The complete data density function is thus:

$$\begin{aligned} f(y_1, y_2, z, \ln(\mu)|\theta) &= p(\ln(\mu), y_2, z|\theta)p(y_1|\ln(\mu)) \\ &= \prod_{i=1}^{N} f(s_{i1}|\theta) f_{Po}(y_{1i}|\ln(\mu_{i1})) \mathbb{1}\{z_i \in \mathcal{B}_i\}) \end{aligned} \tag{19}$$

where $\mathcal{B}_i = (-\infty, 0)$ if $T_i = 0$ and $\mathcal{B}_i = [0, +\infty)$ if $T_i = 1$.

Let's partition Ω as follows:

$$\Omega = \begin{pmatrix} \underset{(2\times 2)}{\Omega_{11}} & \Omega_{21} \\ \Omega_{12} & 1 \end{pmatrix} = \begin{pmatrix} \tilde{\Omega_{11}} & \tilde{\Omega_{12}} \\ \tilde{\Omega_{21}} & \underset{(2\times 2)}{\tilde{\Omega_{22}}} \end{pmatrix}$$

To simplify the notation, we define the following functions:

$$\Omega_{11\cdot 2} = \Omega_{11} - \Omega_{12}\Omega_{21}$$

The estimation part for this model would be the same as discussed in section 3.

References

James H Albert and Siddhartha Chib. Bayesian analysis of binary and polychotomous response data. *Journal of the American statistical Association*, 88(422):669–679, 1993.

James H Cardon and Igal Hendel. Asymmetric information in health insurance: evidence from the national medical expenditure survey. *RAND Journal of Economics*, pages 408–427, 2001.

Siddhartha Chib. Marginal likelihood from the gibbs output. *Journal of the american statistical association*, 90(432):1313–1321, 1995.

Siddhartha Chib, Edward Greenberg, and Rainer Winkelmann. Posterior simulation and bayes factors in panel count data models. *Journal of Econometrics*, 86(1):33–54, 1998.

Siddhartha Chib, Edward Greenberg, and Ivan Jeliazkov. Estimation of semiparametric models in the presence of endogeneity and sample selection. *Journal of Computational and Graphical Statistics*, 18(2):321–348, 2009.

David M Cutler and Richard J Zeckhauser. Adverse selection in health insurance. In *Forum for Health Economics & Policy*, volume 1. De Gruyter, 1998.

James J Heckman. Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, pages 153–161, 1979.

Ivan Jeliazkov and Esther Hee Lee. Mcmc perspectives on simulated likelihood estimation. In *Maximum simulated likelihood methods and applications*. Emerald Group Publishing Limited, 2010.

Michael Keane and Olena Stavrunova. Adverse selection, moral hazard and the demand for medigap insurance. *Journal of Econometrics*, 190(1):62–78, 2016.

Murat K Munkin and Pravin K Trivedi. Bayesian analysis of a self-selection model with multiple outcomes using simulation-based estimation: an application to the demand for healthcare. *Journal of Econometrics*, 114(2):197–220, 2003.

Son Nghiem and Nicholas Graves. Selection bias and moral hazard in the australian private health insurance market: Evidence from the queensland skin cancer database. *Economic Analysis and Policy*, 64:259–265, 2019.

James D Reschovsky, Peter Kemper, and Ha Tu. Does type of health insurance affect health care use and assessments of care among the privately insured? *Health services research*, 35(1 Pt 2):219, 2000.

Christian Ritter and Martin A Tanner. Facilitating the gibbs sampler: the gibbs stopper and the griddy-gibbs sampler. *Journal of the American Statistical Association*, 87(419): 861–868, 1992.

Claudio Sapelli and Bernardita Vial. Self-selection and moral hazard in chilean health insurance. *Journal of health economics*, 22(3):459–476, 2003.

Dan Shane. *Incentives and self-selection: Exploring the relationship between private health insurance and health services utilization.* PhD thesis, Indiana University, 2012.