Bayesian Analysis of a Self-selection Model with Multiple Outcomes

Jieyu Gao

March, 2024

Abstract

We developed a Bayesian treatment model that incorporates self-selection and accommodates multiple outcomes. We discussed the estimation of marginal likelihood for formal model comparison. To validate our approach, we tested the algorithm and model comparison techniques through simulation results. Subsequently, we applied our model to two distinct datasets, enabling us to analyze the influence of insurance on healthcare utilization. Specifically, we estimated the impact of Medigap policies on healthcare expenditure using the 1987 National Medical Expenditure Survey (NMES) data and the impact of different types of private insurance on healthcare utilization using the 1996 Medical Expenditure Panel Survey (MEPS) data. Our analysis revealed weak evidence supporting selection bias in both applications.

1 Introduction

This paper introduces a parametric self-selection Bayesian model featuring a binary treatment and two distinct outcome variables—one count and the other continuous. Allowing for endogenous selection, we are interested in studying the treatment impact on the conditional mean of the outcome variables. By leveraging this model, we conducted an empirical investigation into the influence of private insurance on healthcare expenditures and the frequency of physician office visits across two datasets.

Sample selection arises when the observed sample does not accurately represent the broader population of interest, introducing bias into estimators. Heckman (1979) highlighted the consequential inconsistency and bias resulting from ignoring sample selection and introduced Heckman correction as a remedy. In the context of treatment models, we typically only observe treated outcomes for individuals in the treated group and untreated outcomes for those in the untreated group, potentially leading to sample selection bias if selection decisions are non-random. Addressing this challenge, Chib, Greenberg and Jeliazkov (2009) pioneered a Bayesian model capable of analyzing data afflicted by both sample selection and endogeneity issues.

Many papers discussed possible selection concerns in the health insurance industry. Cutler and Zeckhauser (1998) examined evidence suggestive of adverse selection, shedding light on its implications. Keane and Stavrunova (2016) advanced a model aimed at estimating moral hazard and selection dynamics within the Medigap market, revealing a weak adverse selection alongside notable moral hazard effects, especially among individuals with better health conditions. 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. Nghiem and Graves (2019) endeavored to estimate the impact of moral hazard and selection biases within Australia's private insurance landscape, unearthing evidence indicative of advantageous selection tendencies, wherein risk-averse individuals exhibit a heightened propensity to procure health insurance coverage.

On the other hand, Reschovsky, Kemper and Tu (2000) posited that there exists minimal disparity in hospital utilization across various insurance plan types. Cardon and Hendel (2001) found that there is no discernible evidence of asymmetric information significantly influencing insurance plan choices. Munkin and Trivedi (2003) found weak evidence of self-selection in their model estimating the impact of health insurance on health care utilization.

We expand upon the potential outcome methodology employed in Munkin and Trivedi (2003) by relaxing the assumption of constant treatment effects. This enhancement allows for a more nuanced exploration of treatment effects. Incorporating insights from Bayesian research by Albert and Chib (1993), Chib, Greenberg and Winkelmann (1998) and Chib, Greenberg and Jeliazkov (2009), we use the Markov chain Monte Carlo (MCMC) method to estimate the parameters. Our methodology was applied to two datasets to investigate the influence of public or private health insurance on healthcare expenditure and the frequency of doctor's office visits. Specifically, we analyzed the 1996 Medical Expenditure Panel Survey (MEPS) and the 1987 National Medical Expenditure Survey (MEPS) data to gain insights into these impacts. Additionally, we conducted estimations using three alternative parsimonious models to provide a comprehensive comparative analysis. To facilitate model comparison, we computed the marginal likelihood for each model.

The remainder of the paper is structured as follows. In Section 2, we delineate the model, its corresponding estimation algorithm, and the model comparison technique. In section 3, we present the simulation results. Section 4 discusses the empirical findings. Finally, Section 5 concludes the paper.

2 Model

This section considers a potential outcome framework with self-selection for estimating the treatment effect for multiple outcomes. The binary treatment variable is represented by D_i, where $D_i = 1$ if subject i receives treatment and $D_i = 0$ if subject i is in the control group. For each i in the sample, we denote the potential outcomes for the count and continuous outcome variables as y_{j1i} and y_{j2i} respectively, where $j = 0, 1$ signifies the treatment status. We assume that

$$y_{j1i} \sim \text{Poisson}(\mu_{ji}), \quad j = 0, 1.$$

For each i in the sample, the observed response is expressed as

$$\begin{aligned} y_{1i} &= y_{01i} + (y_{11i} - y_{01i})D_i, \\ y_{2i} &= y_{02i} + (y_{12i} - y_{02i})D_i. \end{aligned}$$

We assume that $D_i = \mathbb{1}\{d_i^* \geq 0\}$, where d_i^* represents a latent variable that determines the values of D_i. Additionally, we let x_{d1i} to denote the covariates for y_{d1i}, x_{d2i} to denote the covariates for y_{d2i}, where $d = 0, 1$. Lastly, we use x_{di} to denote the covariates for d_i^*.

The model can be represented as

$$g_i = X_i\beta + \varepsilon_i, \quad \epsilon_i \sim \Omega \tag{1}$$

where

$$g_i = (d_i^*, \ln(\mu_{0i}), y_{02i}, \ln(\mu_{1i}), y_{12i})', \quad X_i = \begin{pmatrix} x'_{di} & 0 & 0 & 0 & 0 \\ 0 & x'_{01i} & 0 & 0 & 0 \\ 0 & 0 & x'_{02i} & 0 & 0 \\ 0 & 0 & 0 & x'_{11i} & 0 \\ 0 & 0 & 0 & 0 & x'_{12i} \end{pmatrix},$$

$$\beta = (\beta'_d, \beta'_{01}, \beta'_{02}, \beta'_{11}, \beta'_{12})', \quad \varepsilon_i = (\varepsilon_{di}, \varepsilon_{01i}, \varepsilon_{02i}, \varepsilon_{11i}, \varepsilon_{12i})'.$$

If $d_i^* \geq 0$, we observe $(\ln(\mu_{1i}), y_{12i})$; otherwise, we observe $(\ln(\mu_{01}), y_{02i})$. We define the vectors $g_{0i} = (d_i^*, \ln(\mu_{0i}), y_{02i})'$ and $g_{1i} = (d_i^*, \ln(\mu_{1i}), y_{12i})'$. The covariates matrix is given by

$$X_{1i} = \begin{pmatrix} x'_{di} & 0 & 0 \\ 0 & x'_{11i} & 0 \\ 0 & 0 & x'_{12i} \end{pmatrix}, \quad X_{0i} = \begin{pmatrix} x'_{di} & 0 & 0 \\ 0 & x'_{02i} & 0 \\ 0 & 0 & x'_{12i} \end{pmatrix}.$$

The covariance matrix is represented as

$$\Omega = \begin{pmatrix} 1 & \omega_{d01} & \omega_{d02} & \omega_{d11} & \omega_{d12} \\ \omega_{01d} & \omega_{01} & \omega_{012} & \omega_{0111} & \omega_{0112} \\ \omega_{02d} & \omega_{021} & \omega_{02} & \omega_{0211} & \omega_{0212} \\ \omega_{11d} & \omega_{1101} & \omega_{1102} & \omega_{11} & \omega_{112} \\ \omega_{12d} & \omega_{1201} & \omega_{1202} & \omega_{121} & \omega_{12} \end{pmatrix}$$

Due to missing data, we cannot identify Ω_{0111}, Ω_{0112}, Ω_{0211}, and Ω_{0212}. To address this identification issue, we assume that $\Omega_d = 1$.

Thus the covariance matrix that can be identified is expressed as

$$\Omega = \begin{pmatrix} 1 & \omega_{d01} & \omega_{d02} & \omega_{d11} & \omega_{d12} \\ \omega_{01d} & \omega_{01} & \omega_{012} & . & . \\ \omega_{02d} & \omega_{021} & \omega_{02} & . & . \\ \omega_{11d} & . & . & \omega_{11} & \omega_{112} \\ \omega_{12d} & . & . & \omega_{121} & \omega_{12} \end{pmatrix}$$

We introduce the following notation

$$\Omega_0 = \begin{pmatrix} 1 & \omega_{d01} & \omega_{d02} \\ \omega_{01d} & \omega_{01} & \omega_{012} \\ \omega_{02d} & \omega_{021} & \omega_{02} \end{pmatrix} = \begin{pmatrix} 1 & \Omega_{012} \\ \Omega_{021} & \underset{(2\times2)}{\Omega_{011}} \end{pmatrix}, \quad \Omega_1 = \begin{pmatrix} 1 & \omega_{d11} & \omega_{d12} \\ \omega_{d11} & \omega_{11} & \omega_{112} \\ \omega_{12d} & \omega_{121} & \omega_{12} \end{pmatrix} = \begin{pmatrix} 1 & \Omega_{112} \\ \Omega_{121} & \underset{(2\times2)}{\Omega_{111}} \end{pmatrix},$$

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

$$\Omega_{22\cdot0} = \Omega_{011} - \Omega_{021}\Omega_{012}, \quad \Omega_{22\cdot1} = \Omega_{111} - \Omega_{121}\Omega_{112}.$$

Given the unit restriction, we directly handle the quantities $\Omega_{22\cdot 0}$, $\Omega_{22\cdot 1}$, Ω_{012}, and Ω_{112}, from which we subsequently recover Ω_0 and Ω_1 (Munkin and Trivedi, 2003; Chib, Greenberg and Jeliazkov, 2009; Vossmeyer, 2016).

Thus for i in the control group, we have

$$g_{0i} \propto |\Omega_0|^{-1/2} \exp\left\{ -\frac{1}{2}(g_{0i} - X_{0i}J_0\beta)'\Omega_0^{-1}(g_{0i} - X_{0i}J_0\beta) \right\}$$

, and for i in the treated group, we have

$$g_{1i} \propto |\Omega_1|^{-1/2} \exp\left\{ -\frac{1}{2}(g_{1i} - X_{1i}J_1\beta)'\Omega_1^{-1}(g_{1i} - X_{1i}J_1\beta) \right\}$$

The complete data density function is given by

$$\begin{aligned} &f(y_1, y_2, d^*, \ln(\mu)|\beta, \Omega_0, \Omega_1) = f(\ln(\mu), y_2, d^*|\theta) f(y_1|\ln(\mu)) \\ &= \left[\prod_{i:D_i=1} f_N(g_{i1}|\beta, \Omega_1) f(y_{1i}|\ln(\mu_{1i})) \mathbb{1}\{d_i^* \geq 0\}) \right] \left[\prod_{i:D_i=0} f(g_{0i}|\beta, \Omega_0) f(y_{1i}|\ln(\mu_{1i})) \mathbb{1}\{d_i^* < 0\} \right] \end{aligned} \tag{2}$$

The prior distribution for β is specified as $\beta \sim N(b_0, B_0)$. We let $\Omega_{22\cdot 0} \sim IW(r_0, R_0)$, $\Omega_{22\cdot 1} \sim IW(r_1, R_1)$, $\Omega_{012}|\Omega_{22\cdot 0} \sim N(q_0, \Omega_{22\cdot 0})$, and $\Omega_{112}|\Omega_{22\cdot 1} \sim N(q_1, \Omega_{22\cdot 1})$.

In our model, the average treatment effect (ATE) and the average treatment effect on the treated (ATT) for outcome Y_1 and Y_2 are defined as

$$ATE_1 = E(Y_{11i} - Y_{01i}) = E(x'_{11i}\beta_{11} - x'_{01i}\beta_{01}), \quad ATE_2 = E(Y_{12i} - Y_{02i}) = E(x'_{12i}\beta_{12} - x'_{02i}\beta_{02}),$$
$$ATT_1 = E(Y_{11i} - Y_{01i}|D_i = 1) = E(x'_{11i}\beta_{11} - x'_{11i}\beta_{01}),$$
$$ATT_2 = E(Y_{12i} - Y_{02i}|D_i = 1) = E(x'_{12i}\beta_{12} - x'_{12i}\beta_{02}).$$

These quantities can be estimated using the MCMC output.

2.1 Markov chain Monte Carlo (MCMC) Algorithm

In this section, we delve into the Markov chain Monte Carlo (MCMC) simulation algorithm tailored for our model. We derive the conditional distribution for each parameter and provide a comprehensive discussion of the sampling methodology employed.

2.1.1 Sampling β

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

$$b = B\left(B_0^{-1}b_0 + \sum_{i\in N_0} J_0' X_{0i}'\Omega_0^{-1} g_{0i} + \sum_{i\in N_1} J_1' X_{1i}'\Omega_1^{-1} g_{1i}\right),$$

and

$$B = \left(B_0^{-1} + \sum_{i\in N_0} J_0' X_{0i}'\Omega_0^{-1} X_{0i} J_0 + \sum_{i\in N_1} J_1' X_{i1}'\Omega_1^{-1} X_{1i} J_1\right)^{-1}.$$

2.1.2 Sampling d_i^*

We sample $d_i^* \sim TN_{\mathcal{B}_i}(\mu_{dji}, \Omega_{dj})$, where $\mathcal{B}_i = (-\infty, 0)$ if $D_i = 0$, and $\mathcal{B}_i = [0, +\infty)$ if $D_i = 1$, $\mu_{dji} = x_{di}'\beta_d + \Omega_{j12}(\Omega_{j11})^{-1}\left(\begin{pmatrix} \ln(\mu_{ji}) - x_{j1i}'\beta_{j1} \\ y_{j2i} - x_{j2i}'\beta_{j2}\end{pmatrix}\right)$, and $\Omega_{dj} = 1 - \Omega_{j12}(\Omega_{j11})^{-1}\Omega_{j21}$, $j = 0, 1$.

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

The conditional distribution of $\ln(\mu_{ji})$ given y_{j1i}, y_{j2i}, d_i^* and all the parameters is

$$(\ln(\mu_{ji})|y_{j2i}, d_i^*, \beta, \Omega_j) \sim N(\mu_{1|2,ji}, \sigma_{1|2,j})$$

where $\mu_{1|2,ji} = x_{j1i}'\beta_{j1} + \tilde{\Omega}_{j12}(\tilde{\Omega}_{j22})^{-1}\begin{pmatrix} d_i^* - x_{di}'\beta_d \\ y_{j2i} - x_{j2i}'\beta_{j2}\end{pmatrix}$, and $\sigma_{1|2,j} = \omega_{j1} - \tilde{\Omega}_{j12}(\tilde{\Omega}_{j22})^{-1}\tilde{\Omega}_{j21}$, $\tilde{\Omega}_{j12} = (\omega_{j1d}, \omega_{j12})$, and $\tilde{\Omega}_{j22} = \begin{pmatrix} 1 & \omega_{dj2} \\ \omega_{j2d} & \omega_{j2}\end{pmatrix}$, $j = 0, 1$. Thus the posterior distribution for $\ln(\mu_{it})$ is specified as

$$\pi(\ln(\mu_{ji})|y_{j1i}, y_{j2i}, d_i^*, \beta, \Omega_j) \propto f_N(\mu_{1|2,ji}, \sigma_{1|2,j}) f(y_{j1i}|\ln(\mu_{ji})) = f_N(\mu_{1|2,ji}, \sigma_{1|2,j})\frac{\mu^{y_{j1i}}e^{-\mu_{ji}}}{y_{j1i}!}$$

We use a random walk Metropolis-Hasting algorithm to sample the $\ln(\mu_{ji})$. We use $\ln(\mu_{ji})^*$ to denote the proposed value. The proposed density is denoted as $q(\ln(\mu_{ji}), \ln(\mu_{ji})^*) = f_N(\ln(\mu_{ji})^*|\ln(\mu_{ji}), ((\sigma_{1|2,j})^{-1} + y_{j1i}^{-1})^{-1})$. The acceptance rate is thus defined as

$$\begin{aligned}\alpha(\ln(\mu_{ji}), \ln(\mu_{ji})^*) &= \min\left\{\frac{\pi(\ln(\mu_{ji})^*|y_{j1i}, y_{j2i}, \beta, d_i^*, \Omega_j)q(\ln(\mu_{ji}), \ln(\mu_{ji}^*))}{\pi(m_i|y_{i1t}, y_{i2t}, \beta, z_i, \Omega)q(\ln(\mu_{ji})^*, \ln(\mu_{ji}))}, 1\right\} \\ &= \min\{\frac{\pi(\ln(\mu_{ji})^*|y_{j1i}, y_{j2i}, \beta, d_i^*, \Omega_j)}{\pi(\ln(\mu_{ji})|y_{j1i}, y_{j2i}, \beta, d_i^*, \Omega_j)}, 1\}.\end{aligned}$$

We generate a random number p from a uniform distribution $U(0, 1)$. If $p \leq \alpha$, the proposed value is accepted; otherwise, it is rejected.

2.1.4 Sampling Ω_{j21}

The conditional distribution of Ω_{j21}, $j = 0, 1$, is given by

$$\begin{aligned}\pi(\Omega_{j21}|\Omega_{22\cdot j}, \beta, \ln(\mu_{ji}), \ln(y_{j2i}), d_i^*) &\propto \pi(\Omega_{j21}|\Omega_{22\cdot j})\pi\left(\begin{pmatrix} \ln(\mu_{ji}) \\ y_{j2i} \end{pmatrix} |d_i^* \right) \\ &= f_N(q_j, \Omega_{22\cdot j}) \prod_{i:D_i=j} f_N((\ln(\mu_{ji}), y_{j2i})'|\mu_{2|3,ji}, \Omega_{22\cdot j})) \quad (3)\end{aligned}$$

where $\mu_{2|3,ji} = \begin{pmatrix} x'_{j1i}\beta_{j1} \\ x'_{j2i}\beta_{j2} \end{pmatrix} + \Omega_{j21}(d_i^* - x'_{di}\beta_d)$.

Thus the posterior distribution for Ω_{j21} is specified as $\Omega_{j21}|\Omega_{22\cdot j} \sim N([\epsilon_{jd}\epsilon_{jd}(\Omega_{22\cdot j})^{-1} + (\Omega_{22\cdot j})^{-1}]^{-1}[(\Omega_{22\cdot j})^{-1}\epsilon_j^{'}\epsilon_{jd} + (\Omega_{22\cdot j}^t)^{-1}q_j], [\epsilon'_{jd}\epsilon_{jd}(\Omega_{22\cdot j})^{-1} + (\Omega_{22\cdot j})^{-1}]^{-1})$, where $\epsilon_{jd} = d_i^* - X_d\beta_d$, and $\epsilon_j = \big(\ln(\mu_j) - X_{j1}\beta_{j1}, \quad y_{j2} - X_{j2}\beta_{j2}\big)$.

2.1.5 Sampling $\Omega_{22\cdot j}$

We sample $\Omega_{22\cdot j}$, $j = 0, 1$, from the distribution specified as

$$\begin{aligned}&\pi(\Omega_{22\cdot j}|\Omega_{j21}, \beta, \ln(\mu_{ji}), y_{j2i}, d_i^*) \\ \propto&\pi(\Omega_{22\cdot j})\pi(\Omega_{j21}|\Omega_{22\cdot j})\pi\left(\begin{pmatrix} \ln(\mu_{ji}) \\ \ln(y_{j2i}) \end{pmatrix} |d_i^* \right) \\ =&\pi(\Omega_{22\cdot j}) f_N(q_j, \Omega_{22\cdot j}) \prod_{i:D_i=j} f_N\left((\ln(\mu_{ji}), y_{j2i})' \,|\mu_{2|3,ji}, \Omega_{22\cdot j}) \right) \\ \propto&|\Omega_{22\cdot j}|^{\frac{r_j+4+n_j}{2}} \times \\ \exp&\left(-\frac{1}{2} \left(tr\left[\left(R_j + (\Omega_{j21} - q_j)(\Omega_{j21} - q_j)' + \sum_{i:D_i=j} (\epsilon_{ji} - \Omega_{j21}\epsilon_{jdi})(\epsilon_{ji} - \Omega_{j12}\epsilon_{jdi})' \right) (\Omega_{22\cdot j})^{-1} \right] \right) \right)\end{aligned}$$

where n_0 represents the sample size of the control group, and n_1 denotes the sample size of the treated group. Thus, we sample $\Omega_{22\cdot j}$ as

$$\Omega_{22\cdot j} \sim \text{Inverse Wishart}(r_j+1+n_j, R_j+(\Omega_{j21}-q_j)(\Omega_{j21}-q_j)' + \sum_{i:D_i=j} (\epsilon_{ji}-\Omega_{j21}\epsilon_{jdi})(\epsilon_{ji}-\Omega_{j21}\epsilon_{jdi})')$$

2.2 Bayesian Model Comparison

Bayesian model comparison offers a systematic approach for comparing multiple competing models. According to Bayes' formula, the posterior probability for a given model, denoted as $\mathcal{M}_c$, is expressed as

$$\pi(\mathcal{M}_c|y) \propto \pi(\mathcal{M}_c)m(y|\mathcal{M}_c)$$

where $\mathcal{M}_c$ signifies the prior probability of model $\mathcal{M}_c$, while $m(y|\mathcal{M}_c)$ denotes the marginal likelihood.

According to Chib (1995), the marginal likelihood can be estimated at a specific parameter value θ_c^* in the parameter space, as

$$(y|\mathcal{M}_c) = \frac{f(y|\mathcal{M}_c, \theta_c^*)\pi(\theta_c^*|\mathcal{M}_c)}{\pi(\theta_c^*|\mathcal{M}_c, y)}. \tag{4}$$

In our model, $\theta_c = \{\beta, \Omega_{22\cdot 0}, \Omega_{012}, \Omega_{22\cdot 1}, \Omega_{112}\} \equiv \{\psi_1, \psi_2, \psi_3, \psi_4, \psi_5\}$. For simplicity, we will omit the notation $\mathcal{M}_c$ in the subsequent discussions. The likelihood in this model can be decomposed as

$$f(y|\theta_c^*) = \left(\prod_{i:D_i=0} f(y_{02i}|\theta_c^*)f(\ln(\mu_{0i}), d_i^*|y_{02i}, \theta_c)f(y_{01i}|\ln(\mu_{0i}))f(D_i|d_i^*)\right) \times \left(\prod_{i:D_i=1} f(y_{12i}|\theta_c^*)f(\ln(\mu_{1i}), d_i^*|y_{12i}, \theta_c)f(y_{11i}|\ln(\mu_{1i}))f(D_i|d_i^*)\right) \tag{5}$$

where $f(y_{02i}|\theta_c^*)$ and $f(y_{12i}|\theta_c^*)$ are straightforward to compute, while the remaining terms can be estimated using the importance sampling method. We employ a Student-t distribution with 5 degrees of freedom as the proposal distribution.

The posterior ordinate $\pi(\theta_c^*|y)$ can be calculated utilizing the Markov Chain's invariant condition (Ritter and Tanner, 1992; Jeliazkov and Lee, 2010), expressed as

$$\pi(\theta_c^*|y) = E(K(\theta_c, \theta_c^*)|y, \xi) \tag{6}$$

where xi denotes the set of latent data, and $K(\cdot)$ represents the Gibbs transition kernel defined as

$$K(\theta_c, \theta_c^*|y) = \pi_{r=1}^5 \pi(\psi_r^*|y, \{\psi_s^*\}_{s<r}, \{\psi_s\}_{s>r}, \xi)$$

where (ψ, ξ) are obtained in the main MCMC run.

"In this paper, we contrast our primary model, designated as $\mathcal{M}_0$, with several more parsimonious alternatives: $\mathcal{M}_1$, which is the baseline model $\mathcal{M}_0$ with constraints $\Omega_{012} =$

$\Omega_{112} = (0, 0)$, M_2, a model featuring a constant treatment effect, and M_3, an extension of M_2 where $\Omega_{12} = (0, 0)$ is imposed.

It's important to note that M_2 offers a more parsimonious version of M_0. The specific model details for M_2 are provided in Appendix A.

3 Simulation

In this section, we conduct a simulation study to evaluate both the performance of the MCMC algorithm and the effectiveness of the proposed model comparison approach. We randomly generated 10,000 observations from the model $\mathcal{M}_0$ using Equation 1, and estimated the model using the method introduced in the previous session. The parameters are specified as

$$\Omega = \begin{pmatrix} 1 & 0.55 & 0.34 & 0.46 & -0.19 \\ 0.55 & 1.04 & -0.13 & -0.05 & -0.56 \\ 0.34 & -0.13 & 0.32 & 0.30 & 0.17 \\ 0.46 & -0.05 & 0.30 & 0.71 & -0.12 \\ -0.19 & -0.56 & 0.17 & -0.12 & 0.64 \end{pmatrix}$$

$\beta_d = (-0.5, 1, -0.5)'$, $\beta_{01} = (-0.5, 0.5)'$, $\beta_{11} = (0.5, -0.5)'$, $\beta_{02} = (1, -1)'$ and $\beta_{12} = (-1, 1)'$. The covariates are $x_{01i} = [1, \nu_{1i}]$, $x_{11i} = x_{02i} = x_{12i} = x_{01i}$, and $x_{di} = [1, \nu_{1i}, \nu_{2i}]$, where ν_{1i} and ν_{2i} are sampled from two independent standard normal distribution.

We estimated the generated data using 10,000 iterations with a burn-in period of 1000 iterations. We reported the mean and standard deviation of the posterior distribution for the parameters, along with the 95% credible interval. The 95% credible intervals are derived from the converged empirical estimate distribution by extracting the 2.5% and 97.5% quantiles. The estimated parameter values are presented in Tables 1 and 2. Based on the results, model $\mathcal{M}_0$ can accurately estimate the parameters, whereas all other models fail to do so effectively.

The findings regarding treatment effects are summarized in Tables 3 and 4. Additionally, we present the posterior distribution of ATE_1, ATE_2, ATT_1, and ATT_2 in Figures 1 and 2. Notably, Model $\mathcal{M}_0$ demonstrates accurate estimations of the treatment effects compared to alternative models, which exhibit less effective performance in this regard. According to the marginal likelihood results, model $\mathcal{M}_0$ emerges as the favored choice.

Table 1: $\mathcal{M}_0$ and $\mathcal{M}_1$: Estimation Results

		$\mathcal{M}_0$			$\mathcal{M}_1$		
	True	Mean	SD	95% CI	Mean	SD	95% CI
CONST	-0.5	-0.49	0.02	$(-0.52, -0.46)$	-0.49	0.02	$(0.52, -0.46)$
ν_1	1	0.98	0.02	$(0.94, 1.02)$	0.98	0.02	$(0.94, 1.02)$
ν_2	-0.5	-0.49	0.02	$(-0.52, -0.47)$	-0.49	0.02	$(-0.52, -0.46)$
CONST	-0.5	-0.53	0.03	$(-0.59, -0.46)$	-0.83	0.03	$(-0.89, -0.77)$
ν_1	0.5	0.45	0.03	$(0.38, 0.51)$	0.23	0.03	$(0.18, 0.27)$
CONST	1	0.99	0.01	$(0.97, 1.01)$	0.82	0.01	$(0.81, 0.84)$
ν_1	-1	-1.01	0.01	$(-1.03, -0.99)$	-1.11	0.01	$(-1.13, -1.10)$
CONST	0.5	0.48	0.06	$(0.36, 0.63)$	0.91	0.02	$(0.87, 0.95)$
ν_1	-0.5	-0.49	0.04	$(-0.56, -0.42)$	-0.68	0.02	$(-0.73, -0.64)$
CONST	-1	-1.01	0.05	$(-1.11, -0.93)$	-1.18	0.02	$(-1.22, -1.15)$
ν_1	1	1.01	0.03	$(0.96, 1.06)$	1.08	0.02	$(1.05, 1.11)$
ω_{01}	1.04	1.11	0.06	$(0.99, 1.23)$	0.94	0.05	$(0.85, 1.04)$
ω_{012}	-0.13	-0.13	0.01	$(-0.16, -0.11)$	-0.21	0.01	$(-0.23, -0.18)$
ω_{02}	0.32	0.31	0.01	$(0.30, 0.33)$	0.28	0.00	$(0.27, 0.29)$
ω_{d01}	0.55	0.58	0.05	$(0.48, 0.67)$			
ω_{d02}	0.34	0.32	0.02	$(0.28, 0.35)$			
ω_{11}	0.71	0.68	0.04	$(0.60, 0.75)$	0.59	0.03	$(0.54, 0.64)$
ω_{112}	-0.12	-0.12	0.02	$(-0.16, -0.09)$	-0.09	0.01	$(-0.12, -0.06)$
ω_{12}	0.65	0.64	0.02	$(0.61, 0.67)$	0.63	0.01	$(0.60, 0.65)$
ω_{d11}	0.46	0.43	0.06	$(0.29, 0.55)$			
ω_{d12}	-0.19	-0.17	0.04	$(-0.25, -0.08)$			

4 Empirical Application

We apply our model to analyze two datasets to investigate the impact of insurance on healthcare utilization. The first dataset is sourced from the 1987-1988 National Medical Expenditure Survey (NMES), encompassing data on the U.S. elderly population with positive medical expenditures. The second dataset is derived from the 1996 Medical Expenditure Panel Survey (MEPS), comprising non-elderly individuals privately insured with positive medical expenditures. Table 5 provides a summary of statistics and variable definitions for both datasets.

Table 2: $\mathcal{M}_2$ and $\mathcal{M}_3$: Estimation Results

	$\mathcal{M}_2$			$\mathcal{M}_3$		
	Mean	SD	95% CI	Mean	SD	95% CI
CONST	-0.34	0.01	$(-0.37, -0.32)$	-0.34	0.01	$(-0.37, -0.32)$
ν_1	-0.01	0.01	$(-0.03, 0.01)$	-0.01	0.01	$(-0.03, 0.02)$
ν_2	-0.03	0.01	$(-0.05, -0.01)$	-0.03	0.01	$(-0.06, -0.01)$
CONST	-1.43	0.07	$(-1.55, -1.29)$	-0.88	0.02	$(-0.93, -0.83)$
ν_1	-0.14	0.02	$(-0.18, -0.10)$	-0.12	0.02	$(-0.15, -0.08)$
D	2.82	0.16	$(2.47, 3.10)$	1.37	0.03	$(1.31, 1.44)$
CONST	1.25	0.04	$(1.17, 1.33)$	1.13	0.01	$(1.11, 1.16)$
ν_1	-0.33	0.01	$(-0.35, -0.30)$	-0.33	0.01	$(-0.36, -0.31)$
D	-1.67	0.11	$(-1.90, -1.46)$	-1.37	0.03	$(-1.42, -1.32)$
ω_1	1.33	0.10	$(1.14, 1.53)$	0.87	0.03	$(0.82, 0.93)$
ω_{12}	-0.56	0.05	$(-0.66, -0.48)$	-0.46	0.02	$(-0.50, -0.43)$
ω_2	1.24	0.02	$(1.20, 1.29)$	1.22	0.02	$(1.18, 1.25)$
ω_{d1}	-0.85	0.09	$(-1.00, -0.66)$			
ω_{d2}	0.19	0.07	$(-0.06, 0.32)$			

Table 3: $\mathcal{M}_0$ and $\mathcal{M}_1$: Treatment Effects

		$\mathcal{M}_0$			$\mathcal{M}_1$		
	True	Mean	SD	95% CI	Mean	SD	95% CI
ATE_1	1.01	1.02	0.07	$(0.88, 1.17)$	1.74	0.04	$(1.67, 1.82)$
ATT_1	0.33	0.38	0.07	$(0.25, 0.52)$	1.13	0.04	$(1.05, 1.21)$
ATE_2	-2.01	-2.02	0.05	$(2.12, -1.93)$	-2.02	0.02	$(-2.06, -1.98)$
ATT_2	-0.66	-0.65	0.04	$(-0.73, -0.58)$	-0.53	0.02	$(-0.57, -0.50)$
$\ln(m(y))$				-27622.50			-27803.92

4.1 Private Insurance

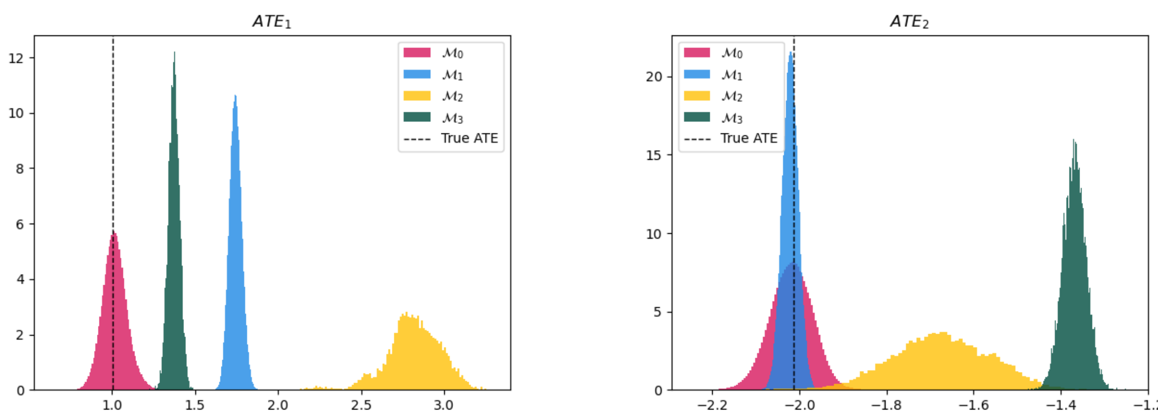
Our first application aims to examine the influence of private insurance on the number of physician doctor office-based visits and the corresponding office-based total expenditures. We collected data from 3680 observations of elderly Americans from the National Medical Expenditure Survey (NMES) conducted in 1987.

Individuals aged 65 and older are typically covered by Medicare, which caters to a broad spectrum of healthcare services. Some individuals opt to supplement this coverage with

Table 4: $\mathcal{M}_2$ and $\mathcal{M}_3$: Treatment Effects

		$\mathcal{M}_2$			$\mathcal{M}_3$		
	True	Mean	SD	95% CI	Mean	SD	95% CI
ATE_1	1.01	2.82	0.16	$(2.47, 3.10)$	1.37	0.03	$(1.31, 1.44)$
ATT_1	0.33						
ATE_2	-2.01	-1.67	0.11	$(-1.90, -1.46)$	-1.37	0.03	$(-1.42, -1.32)$
ATT_2	-0.66						
$\ln(m(y))$				-35832.09			-35832.91

Figure 1: ATE Histograms

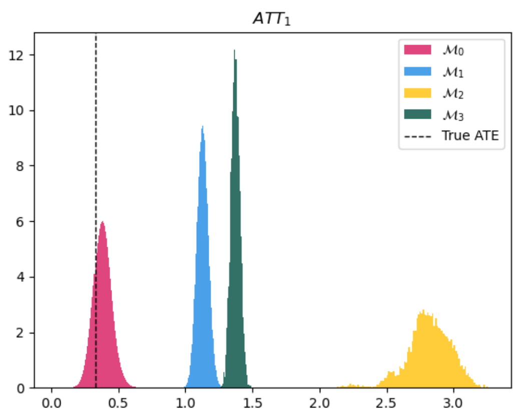


(a) ATE_1 (b) ATE_2

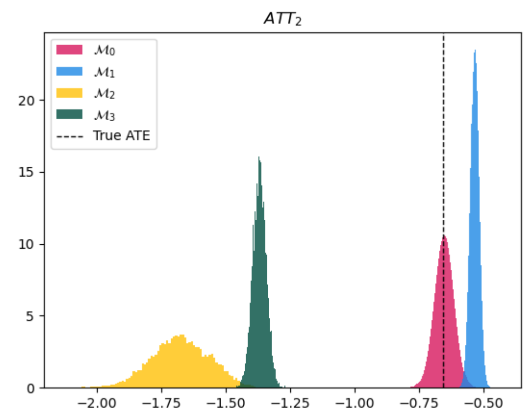
private insurance, particularly if they have chronic health conditions or poor health status.

We apply our model to assess the influence of private insurance on the number of physician office-based visits and associated expenditures. We incorporate covariates such as self-perceived health status, number of chronic conditions, disability status, geographic location, demographic factors, and insurance variables into the equations determining the number of physician office visits and healthcare expenditures. While there may be heterogeneous effects attributable to various types of private insurance policies, our primary focus is on investigating the overall impact of Medigap plans. We assume that family income influences the purchase of private insurance but does not directly affect healthcare utilization. Additionally, we assume that Medicaid has no impact on the selection equation once family

Figure 2: ATE Histograms



(a) ATT_1



(b) ATT_2

income has been controlled for.

We report the posterior mean and standard deviation for the parameter estimated. The results are summarized in Table 6, 7, 8, and 9. Coefficients significantly different from zero at the 95% confidence level are highlighted using asterisks. It's worth noting that the standard deviation for the control group is larger than that obtained for the treatment group. This discrepancy arises from the fact that only 20% of the sample comprises the control group. Due to the larger standard deviations in the control group, certain coefficients that are not substantially different from zero in the control group appear significant in the treated group. These include variables such as $EXCHLTH$, $NORTHEAST$, and $SCHOOL$ in both the doctor's visit and expenditure equations. Conversely, estimates for coefficients of other variables, such as $ADLDIFF$ and $MEDICAID$, are significant in the control group but not in the treated group.

We estimated the treatment effects and marginal likelihood using the four models, with the results summarized in Tables 10 and 11. The treatment effects estimated are not significant in the unrestricted models $\mathcal{M}_0$ and $\mathcal{M}_2$, under the endogeneity assumption. However, the treatment effects estimated are significantly positive for the restricted models $\mathcal{M}_1$ and $\mathcal{M}_3$. This suggests that after accounting for the correlation between insurance and utiliza-

Table 5: Variable Definition and Summary Statistics

Variable	Data set Number of Observations	NMES 3680		MEPS 5368	
	Definition	Mean	SD	Mean	SD
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	size of the company			127.5	177.8
LOCATION	=1 if multiple locations			0.52	0.50
GOVT	= 1 if the company is governmental			0.18	0.39

tion, private insurance has no significant impact on healthcare utilization, indicating the presence of selection bias. We observed significant estimates for the covariance components ω_{d11} and ω_{d12} for the treated group in model $\mathcal{M}0$ to support the statement. However, the estimates for $\omega d01$ and ω_{d02} in $\mathcal{M}0$, as well as the estimates for $\omega d1$ and ω_{d2} in model $\mathcal{M}_2$, are not significant. Thus, we remain uncertain about the presence of selection bias. Based on the marginal likelihood results, model $\mathcal{M}_3$ is favored over the other competing models. Notably, this model suggests a positive impact of Medigap policies on healthcare utilization, indicating the presence of a moral hazard issue.

Table 6: Private Insurance (NMES): MCMC Parameters Estimates for $\mathcal{M}_0$

	INS		VIS(0)		EXP(0)		VIS(1)		EXP(1)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
CONST	−0.20	0.34	1.69*	0.42	4.71*	0.60	1.11*	0.23	4.15*	0.31
EXCHLTH	0.07	0.10	0.02	0.15	−0.01	0.20	−0.19*	0.06	−0.18*	0.07
POORHLTH	−0.05	0.07	0.32*	0.08	0.35*	0.12	0.33*	0.05	0.37*	0.07
NUMCHRON	−0.03	0.02	0.14*	0.02	0.12*	0.03	0.14*	0.01	0.13*	0.02
ADLDIFF	−0.20*	0.07	0.16*	0.08	0.32*	0.12	0.07	0.04	−0.01	0.06
NORTHEAST	0.12	0.07	0.07	0.10	0.10	0.13	0.10*	0.04	0.27*	0.06
MIDWEST	0.27*	0.07	−0.08	0.10	−0.23	0.15	0.03	0.04	0.03	0.06
WEST	−0.15*	0.07	0.22*	0.10	0.50*	0.14	0.06	0.05	0.27*	0.06
BLACK	−0.80*	0.07	0.09	0.12	0.30	0.24	−0.19*	0.07	−0.45*	0.10
MALE	−0.01	0.06	−0.01	0.08	−0.08	0.11	−0.02	0.03	0.01	0.05
MARRIED	0.25*	0.06	−0.07	0.09	−0.14	0.13	−0.01	0.04	0.07	0.05
SCHOOL	0.11*	0.01	−0.01	0.02	−0.03	0.03	0.03*	0.01	0.06*	0.01
AGE	−0.01	0.04	−0.11*	0.05	−0.09	0.07	−0.02	0.03	0.00	0.04
EMPLOYED	0.10	0.09	0.08	0.14	−0.04	0.20	−0.02	0.05	−0.01	0.07
MEDICAID			0.17*	0.08	0.21*	0.10	0.20	0.12	0.20	0.16
FAMINC	0.00*	0.00								
ω_{j1}			0.59*	0.07			0.54*	0.02		
ω_{j12}			0.85*	0.11			0.75*	0.03		
ω_{j2}			1.70*	0.30			1.40*	0.05		
ω_{dj1}			−0.20	0.17			0.15*	0.05		
ω_{dj2}			−0.64	0.39			0.70*	0.07		

4.2 FFS or HMO

In this application, we compare the impact of selecting different types of private insurance on healthcare utilization using data from the 1996 Medical Expenditure Panel Survey (MEPS). The sample comprises individuals aged between 16 and 65. We are comparing two categories of private insurance: indemnity plans (FFS) and HMO plans. FFS is a payment system where patients pay for services upfront and are later reimbursed by the insurance company.

Table 7: Private Insurance (NMES): MCMC Parameters Estimates for $\mathcal{M}_1$

	INS		VIS(0)		EXP(0)		VIS(1)		EXP(1)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
CONST	−0.24	0.35	1.81*	0.40	5.09*	0.53	1.24*	0.22	4.64*	0.30
EXCHLTH	0.06	0.10	0.03	0.14	0.02	0.19	−0.19*	0.05	−0.19*	0.07
POORHLTH	−0.06	0.07	0.31*	0.08	0.33*	0.11	0.33*	0.05	0.41*	0.07
NUMCHRON	−0.04	0.02	0.13*	0.02	0.10*	0.03	0.14*	0.01	0.14*	0.02
ADLDIFF	−0.24*	0.07	0.14	0.08	0.23*	0.10	0.09*	0.04	0.07	0.06
NORTHEAST	0.15*	0.07	0.08	0.09	0.16	0.12	0.09	0.04	0.23*	0.06
MIDWEST	0.29*	0.07	−0.04	0.09	−0.10	0.13	0.01	0.04	−0.04	0.05
WEST	−0.17*	0.07	0.20*	0.09	0.43*	0.12	0.07	0.05	0.30*	0.06
BLACK	−0.85*	0.08	−0.02	0.07	−0.06	0.10	−0.11	0.07	−0.14	0.09
MALE	−0.04	0.06	−0.01	0.08	−0.09	0.11	−0.02	0.04	0.02	0.05
MARRIED	0.27*	0.06	−0.03	0.08	0.00	0.10	−0.03	0.04	−0.01	0.05
SCHOOL	0.11*	0.01	0.01	0.01	0.02	0.01	0.02*	0.01	0.03*	0.01
AGE	0.00	0.04	−0.11*	0.05	−0.08	0.07	−0.02	0.03	0.00	0.04
EMPLOYED	0.09	0.10	0.10	0.14	0.04	0.19	−0.02	0.05	−0.02	0.07
MEDICAID			0.16*	0.07	0.20*	0.10	0.20	0.12	0.23	0.16
FAMINC	0.01*	0.00								
ω_1			0.54*	0.04			0.53*	0.02		
ω_{12}			0.73*	0.05			0.71*	0.02		
ω_2			1.31*	0.07			1.24*	0.03		

This arrangement can potentially lead to overuse of healthcare services, as doctors may be incentivized to prescribe more treatments or services.

The treatment variable $INSURANCE$ equals 1 if the insurance type is indemnity plans (FFS) and 0 if it's HMO. Similar to the previous section, we assume that employment-related variables, such as company size, location, self-employment status, and family income, solely influence the selection of private insurance and do not directly impact healthcare utilization. This assumption is grounded in reality, as many individuals have limited options for health insurance plans, often tied closely to their employment circumstances.

The posterior mean and standard deviation for the four models are summarized in Tables 12, 13, 14, and 15. Coefficients significant at the 95% level are highlighted with asterisks. According to Table 12, the estimated covariance between insurance selection and healthcare

Table 8: Private Insurance (NMES): MCMC Parameters Estimates for $\mathcal{M}_2$

	INS		VIS		EXP	
	Mean	SD	Mean	SD	Mean	SD
CONST	-0.23	0.35	1.30^*	0.22	4.58^*	0.29
EXCHLTH	0.06	0.10	-0.17^*	0.05	-0.17^*	0.07
POORHLTH	-0.06	0.08	0.32^*	0.04	0.38^*	0.06
NUMCHRON	-0.04	0.02	0.13^*	0.01	0.13^*	0.01
ADLDIFF	-0.24^*	0.07	0.09^*	0.04	0.11^*	0.05
NOREAST	0.14^*	0.07	0.09^*	0.04	0.22^*	0.05
MIDWEST	0.29^*	0.07	0.01	0.04	-0.04	0.05
WEST	-0.17^*	0.07	0.09^*	0.04	0.32^*	0.05
BLACK	-0.84^*	0.08	-0.09	0.07	-0.12	0.09
MALE	-0.04	0.06	-0.01	0.03	0.00	0.04
MARRIED	0.26^*	0.06	-0.03	0.04	-0.01	0.05
SCHOOL	0.11^*	0.01	0.02^*	0.01	0.03^*	0.01
AGE	0.00	0.04	-0.04	0.03	-0.02	0.03
EMPLOYED	0.08	0.10	-0.02	0.05	-0.02	0.06
MEDICAID			0.18^*	0.06	0.23^*	0.08
FAMINC	0.01^*	0.00				
PRIVATE			0.11	0.20	0.26	0.23
ω_1	0.53^*	0.02				
ω_{12}	0.72^*	0.02				
ω_2	1.26^*	0.03				
ω_{d1}	0.04	0.11				
ω_{d2}	0.03	0.13				

utilization is significantly different from zero in model $\mathcal{M}_0$. Specifically, the covariance is negative in the control group and positive in the treated group. However, the estimated covariance between insurance selection and healthcare utilization is not significant in model $\mathcal{M}_2$, as indicated in Table 14.

The estimated treatment effects and marginal likelihood are summarized in Tables 16 and 17. Model $\mathcal{M}_0$ predicted a positive treatment effect of the FFS plan on the frequency of physician office visits, while the impact on healthcare out-of-pocket expenditure remains unclear. However, the estimated treatment effects using all the other models are not significant. Based on the results using $\mathcal{M}_1$, after restricting that the correlation between insurance

Table 9: Private Insurance (NMES): MCMC Parameters Estimates for $\mathcal{M}_3$

	INS		VIS		EXP	
	Mean	SD	Mean	SD	Mean	SD
CONST	-0.24	0.36	1.27^*	0.20	4.55^*	0.26
EXCHLTH	0.06	0.10	-0.17^*	0.05	-0.16^*	0.07
POORHLTH	-0.06	0.07	0.32^*	0.04	0.38^*	0.06
NUMCHRON	-0.04	0.02	0.14^*	0.01	0.13^*	0.01
ADLDIFF	-0.24^*	0.07	0.10^*	0.04	0.11^*	0.05
NOREAST	0.14	0.07	0.08^*	0.04	0.21^*	0.05
MIDWEST	0.29^*	0.07	0.00	0.04	-0.05	0.05
WEST	-0.17^*	0.07	0.09^*	0.04	0.33^*	0.06
BLACK	-0.85^*	0.08	-0.07	0.05	-0.10	0.07
MALE	-0.04	0.06	-0.01	0.03	0.00	0.04
MARRIED	0.26^*	0.06	-0.03	0.03	-0.01	0.04
SCHOOL	0.11^*	0.01	0.02^*	0.00	0.03^*	0.01
AGE	0.00	0.04	-0.04	0.02	-0.02	0.03
EMPLOYED	0.08	0.10	-0.02	0.05	-0.02	0.06
MEDICAID			0.19^*	0.06	0.23^*	0.08
FAMINC	0.01^*	0.00				
PRIVATE			0.18^*	0.04	0.32^*	0.06
ω_1	0.53^*	0.02				
ω_{12}	0.71^*	0.02				
ω_2	1.26^*	0.03				

Table 10: Private Insurance (NMES): Treatment Effects $\mathcal{M}_0$ and $\mathcal{M}_1$

	$\mathcal{M}_0$			$\mathcal{M}_1$		
	Mean	SD	95% CI	Mean	SD	95% CI
ATE_1	0.45	0.26	$(-0.09, 0.97)$	0.20	0.05	$(0.10, 0.29)$
ATT_1	0.49	0.27	$(-0.09, 1.04)$	0.20	0.05	$(0.10, 0.31)$
ATE_2	1.06	0.59	$(-0.23, 1.88)$	0.32	0.06	$(0.20, 0.44)$
ATT_2	1.17	0.63	$(-0.20, 2.04)$	0.32	0.07	$(0.19, 0.45)$
$\ln(m(y))$			-16176.61			-16248.05

selection and healthcare utilization is 0, the average treatment effects estimates are no longer significant. This suggests some evidence of selection bias. However, the difference between models $\mathcal{M}_2$ and $\mathcal{M}_3$ is not significant. Thus, the presence of selection bias remains uncer-

Table 11: Private Insurance (NMES): Treatment Effects $\mathcal{M}_2$ and $\mathcal{M}_3$

	$\mathcal{M}_2$			$\mathcal{M}_3$		
	Mean	SD	95% CI	Mean	SD	95% CI
ATE_1	0.11	0.20	$(-0.28, 0.47)$	0.18	0.04	$(0.10, 0.27)$
ATE_2	0.26	0.23	$(-0.20, 0.68)$	0.32	0.06	$(0.20, 0.43)$
$\ln(m(y))$			-16156.89			-16147.03

tain. According to the marginal likelihood results, the constant treatment model, $\mathcal{M}_2$, is recommended.

5 Conclusion

In conclusion, we introduced a parametric self-selection model with multiple outcomes within a Bayesian framework to estimate the impact of health insurance on healthcare expenditures. Our model incorporates two outcomes: the number of doctor's office visits and healthcare expenditure. We developed a simulation-based estimation algorithm for parameter estimation. Additionally, we proposed an approach to estimate the marginal likelihood for Bayesian model comparison.

Simulation studies have been conducted to evaluate the performance of the MCMC algorithm and the model comparison techniques. Based on the results, our algorithm can accurately estimate the true parameters, and the model comparison results corroborate the true model.

We employed this model in two empirical applications and compared the results with those obtained using three other, more parsimonious models. Our findings provide some weak evidence supporting the presence of selection bias in both applications.

Table 12: FFS vs. HMO (MEPS): MCMC Parameters Estimates for $\mathcal{M}_0$

	INS		VIS(0)		EXP(0)		VIS(1)		EXP(1)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
CONST	0.01	0.12	−0.12	0.14	3.80*	0.19	0.28*	0.13	3.38*	0.17
EXCHLTH	0.10	0.04	−0.28*	0.04	−0.18*	0.05	−0.17*	0.04	0.00	0.05
POORHLTH	−0.04	0.12	0.37*	0.12	0.43*	0.17	0.23*	0.11	0.32	0.16
NUMCHRON	0.00	0.02	0.21*	0.02	0.22*	0.02	0.23*	0.01	0.27*	0.02
INJURY	0.01	0.02	0.17*	0.02	0.24*	0.03	0.12*	0.02	0.11*	0.03
BLACK	−0.21*	0.06	−0.08	0.06	−0.18*	0.08	−0.16*	0.07	−0.32*	0.09
FEMALE	−0.07	0.04	0.39*	0.04	0.38*	0.05	0.30*	0.04	0.33*	0.05
MARRIED	0.05	0.04	−0.07	0.04	−0.10*	0.05	0.00	0.04	0.01	0.06
SCHOOL	0.00	0.01	0.03*	0.01	0.04*	0.01	0.02*	0.01	0.03*	0.01
EMPLOYED	−0.16*	0.05	0.03	0.05	0.02	0.07	−0.14*	0.05	−0.18*	0.06
AGE	0.08*	0.02	−0.01	0.02	0.06*	0.02	0.05*	0.02	0.14*	0.02
NORTHEAST	−0.11*	0.05	0.14*	0.05	0.02	0.07	0.09	0.05	0.07	0.07
MIDWEST	0.12*	0.05	−0.10	0.05	−0.25*	0.07	−0.03	0.04	−0.06	0.06
WEST	−0.40*	0.05	0.09	0.05	0.13	0.07	−0.09	0.06	−0.29*	0.07
SIZE	−0.00*	0.00								
GOVT	0.04	0.04								
LOCATION	−0.03	0.04								
SELFEMP	0.05	0.06								
FAMINC	−0.00*	0.00								
ω_{j1}			0.83*	0.06			0.59*	0.03		
ω_{j12}			1.01*	0.08			0.87*	0.05		
ω_{j2}			1.60*	0.10			1.89*	0.10		
ω_{dj1}			−0.72*	0.06			0.20*	0.05		
ω_{dj2}			−0.62*	0.11			0.97*	0.07		

Table 13: FFS vs. HMO (MEPS): MCMC Parameters Estimates for $\mathcal{M}_0$

	INS		VIS(0)		EXP(0)		VIS(1)		EXP(1)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
CONST	0.00	0.12	0.47*	0.12	4.30*	0.16	0.45*	0.12	4.14*	0.15
EXCHLTH	0.11*	0.04	−0.23*	0.04	−0.14*	0.05	−0.18*	0.04	−0.07	0.05
POORHLTH	0.01	0.12	0.42*	0.11	0.47*	0.16	0.24*	0.10	0.35*	0.14
NUMCHRON	0.01	0.02	0.22*	0.02	0.22*	0.02	0.23*	0.01	0.27*	0.02
INJURY	0.01	0.02	0.18*	0.02	0.25*	0.03	0.12*	0.02	0.11*	0.03
BLACK	−0.21*	0.06	−0.20*	0.06	−0.28*	0.07	−0.12	0.07	−0.17*	0.08
FEMALE	−0.07	0.04	0.36*	0.04	0.35*	0.05	0.31*	0.04	0.37*	0.05
MARRIED	0.02	0.04	−0.06	0.04	−0.10	0.05	0.00	0.04	0.00	0.05
SCHOOL	0.00	0.01	0.03*	0.01	0.04*	0.01	0.02*	0.01	0.04*	0.01
EMPLOYED	−0.14*	0.06	−0.08	0.05	−0.08	0.07	−0.10*	0.04	−0.03	0.06
AGE	0.08*	0.02	0.02	0.02	0.08*	0.02	0.04*	0.02	0.08*	0.02
NORTHEAST	−0.12*	0.05	0.09	0.05	−0.03	0.06	0.11*	0.05	0.15*	0.06
MIDWEST	0.15*	0.05	−0.04	0.05	−0.20*	0.06	−0.05	0.04	−0.16*	0.06
WEST	−0.43*	0.05	−0.07	0.05	−0.03	0.06	−0.03	0.05	−0.03	0.07
SIZE	−0.00*	0.00								
GOVT	−0.02	0.05								
LOCATION	−0.05	0.05								
SELFEMP	0.12	0.07								
FAMINC	0.00	0.00								
ω_{j1}			0.53*	0.02			0.54*	0.02		
ω_{j12}			0.74*	0.03			0.74*	0.02		
ω_{j2}			1.35*	0.04			1.32*	0.04		

Table 14: FFS vs. HMO (MEPS): MCMC Parameters Estimates for $\mathcal{M}_2$

	INS		VIS		EXP	
	Mean	SD	Mean	SD	Mean	SD
CONST	0.00	0.12	0.49*	0.16	4.34*	0.22
EXCHLTH	0.11*	0.04	−0.20*	0.03	−0.09*	0.04
POORHLTH	0.00	0.12	0.33*	0.08	0.40*	0.11
NUMCHRON	0.01	0.02	0.22*	0.01	0.25*	0.02
INJURY	0.01	0.02	0.14*	0.01	0.17*	0.02
BLACK	−0.22*	0.06	−0.17*	0.05	−0.26*	0.07
FEMALE	−0.07	0.04	0.33*	0.03	0.36*	0.03
MARRIED	0.03*	0.04	−0.03	0.03	−0.05	0.04
SCHOOL	0.00	0.01	0.03*	0.00	0.04*	0.01
EMPLOYED	−0.14*	0.06	−0.10*	0.04	−0.07	0.06
AGE	0.08*	0.02	0.03*	0.01	0.09*	0.02
NORTHEAST	−0.12*	0.05	0.10*	0.04	0.05	0.05
MIDWEST	0.15*	0.05	−0.04	0.04	−0.16*	0.05
WEST	−0.42*	0.05	−0.06	0.06	−0.06	0.08
SIZE	−0.00*	0.00				
GOVT	−0.01	0.05				
LOCATION	−0.06	0.05				
SELFEMP	0.11	0.07				
FAMINC	−0.00	0.00				
INSURANCE			−0.07	0.25	−0.25	0.38
ω_1	0.56*	0.03				
ω_{12}	0.76*	0.03				
ω_2	1.39*	0.06				
ω_{d1}	0.04	0.16				
ω_{d2}	0.18	0.24				

Table 15: FFS vs. HMO (MEPS): MCMC Parameters Estimates for $\mathcal{M}_3$

	INS		VIS		EXP	
	Mean	SD	Mean	SD	Mean	SD
CONST	0.00	0.12	0.46^*	0.09	4.19^*	0.11
EXCHLTH	0.11^*	0.04	-0.21^*	0.03	-0.10^*	0.04
POORHLTH	0.01	0.12	0.33^*	0.08	0.41^*	0.11
NUMCHRON	0.01	0.02	0.22^*	0.01	0.25^*	0.02
INJURY	0.01	0.02	0.14^*	0.01	0.16^*	0.02
BLACK	-0.21^*	0.06	-0.16^*	0.04	-0.23^*	0.06
FEMALE	-0.07	0.04	0.33^*	0.03	0.36^*	0.03
MARRIED	0.02	0.04	-0.03	0.03	-0.05	0.04
SCHOOL	0.00	0.01	0.03^*	0.00	0.04^*	0.01
EMPLOYED	-0.15^*	0.06	-0.09^*	0.03	-0.05	0.04
AGE	0.08^*	0.02	0.03^*	0.01	0.09^*	0.01
NORTHEAST	-0.12^*	0.05	0.10^*	0.03	0.06	0.04
MIDWEST	0.15^*	0.05	-0.05	0.03	-0.18^*	0.04
WEST	-0.43^*	0.05	-0.06	0.04	-0.02	0.05
SIZE	-0.00^*	0.00				
GOVT	-0.02	0.05				
LOCATION	-0.05	0.05				
SELFEMP	0.12	0.07				
FAMINC	-0.00	0.00				
INSURANCE			-0.00	0.02	0.03	0.03
omega_11	0.54	0.01				
omega_12	0.74	0.02				
omega_22	1.34	0.03				

Table 16: FFS vs. HMO (MEPS): Treatment Effects $\mathcal{M}_0$ and $\mathcal{M}_1$

	$\mathcal{M}_0$			$\mathcal{M}_1$		
	Mean	SD	95% CI	Mean	SD	95% CI
ATE_1	0.42	0.07	$(0.28, 0.55)$	-0.01	0.03	$(-0.06, 0.04)$
ATT_1	0.45	0.07	$(0.31, 0.58)$	-0.01	0.03	$(-0.06, 0.05)$
ATE_2	-0.22	0.11	$(-0.44, -0.01)$	0.03	0.03	$(-0.03, 0.10)$
ATT_2	-0.17	0.11	$(-0.39, 0.05)$	0.03	0.03	$(-0.03, 0.10)$
$\ln(m(y))$			-23497.00			-23586.72

Table 17: FFS vs. HMO (MEPS): Treatment Effects $\mathcal{M}_2$ and $\mathcal{M}_3$

	$\mathcal{M}_2$			$\mathcal{M}_3$		
	Mean	SD	95% CI	Mean	SD	95% CI
ATE_1	-0.07	0.25	$(-0.60, 0.42)$	0.00	0.02	$(-0.05, 0.04)$
ATE_2	-0.25	0.38	$(-0.84, 0.63)$	0.03	0.03	$(-0.03, 0.10)$
$\ln(m(y))$			-23493.22			-23497.67

Appendix A M_2: Model Specification

In this section, we outline the model for $\mathcal{M}2$. Assume we have a sample comprising of n independent observations, where $y1i$ and y_{2i} represent the outcome variables, where y_{1i} denotes the discrete potential outcome variable, while y_{2i} represents the potential continuous outcome variable. Let D_i denote the treatment status for individual i, with $D_i = \mathbb{1}\{d_i^* \geq 0\}$, where d_i^* represents a latent variable. Furthermore, let x_{1i} denote the covariates for y_{1i}, x_{2i} denote the covariates for y_{2i}, and x_{di} represent the covariates that determines d_i^* for individual i. We assume that

$$y_{1i} \sim \text{Poisson}(\mu_i) \tag{7}$$

The model can be represented as

$$g_i = X_i\beta + \varepsilon_i$$

where

$$g_i = (d_i^*, \ln(\mu_i), y_{2i})', \quad X_i = \begin{pmatrix} x'_{di} & 0 & 0 \\ 0 & x'_{1i} & 0 \\ 0 & 0 & x'_{2i} \end{pmatrix}$$

$\beta = (\beta'_d, \beta'_1, \beta'_2)'$ and $\varepsilon_i = (\varepsilon_{di}, \varepsilon_{1i}, \varepsilon_{2i})'$. Assume that $\varepsilon_i \sim N(0, \Omega)$, where

$$\Omega = \begin{pmatrix} 1 & \omega_{d1} & \omega_{d2} \\ \omega_{1d} & \omega_1 & \omega_{12} \\ \omega_{2d} & \omega_{21} & \omega_2 \end{pmatrix} = \begin{pmatrix} 1 & \Omega_{12} \\ \Omega_{21} & \underset{(2\times 2)}{\Omega_{11}} \end{pmatrix}$$

The complete data density function is expressed as

$$f(y_1, y_2, D, \ln(\mu)|\beta, \Omega) = \prod_{i=1}^{n} f(g_i|\beta, \Omega) f(y_{1i}|\ln(\mu_{1i}))\mathbb{1}\{d_i^* \in \mathcal{B}_i\}) \tag{8}$$

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

Define

$$\Omega_{22} = \Omega_{11} - \Omega_{21}\Omega_{12}.$$

Because of the unit restriction, direct sampling of Ω isn't feasible. As detailed in Section 2, we instead sample Ω_{22} and Ω_{12} and subsequently reconstruct the covariance matrix Ω accordingly. The estimation algorithm for this model closely resembles the one presented in

Section 2.1 and the model proposed by Munkin and Trivedi (2003), hence its omission in this section.

References

Albert, James H, and Siddhartha Chib. 1993. "Bayesian Analysis of Binary and Polychotomous Response Data." *Journal of the American Statistical Association*, 88(422): 669–679.

Cardon, James H, and Igal Hendel. 2001. "Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey." *RAND Journal of Economics*, 408–427.

Chib, Siddhartha. 1995. "Marginal Likelihood from the Gibbs Output." *Journal of the American Statistical Association*, 90(432): 1313–1321.

Chib, Siddhartha, Edward Greenberg, and Ivan Jeliazkov. 2009. "Estimation of Semiparametric Models in the Presence of Endogeneity and Sample Selection." *Journal of Computational and Graphical Statistics*, 18(2): 321–348.

Chib, Siddhartha, Edward Greenberg, and Rainer Winkelmann. 1998. "Posterior Simulation and Bayes Factors in Panel Count Data Models." *Journal of Econometrics*, 86(1): 33–54.

Cutler, David M, and Richard J Zeckhauser. 1998. "Adverse Selection in Health Insurance." Vol. 1, De Gruyter.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica: Journal of the Econometric Society*, 153–161.

Jeliazkov, Ivan, and Esther Hee Lee. 2010. "MCMC Perspectives on Simulated Likelihood Estimation." In *Maximum Simulated Likelihood Methods and Applications*. Emerald Group Publishing Limited.

Keane, Michael, and Olena Stavrunova. 2016. "Adverse Selection, Moral Hazard and the Demand for Medigap Insurance." *Journal of Econometrics*, 190(1): 62–78.

Munkin, Murat K, and Pravin K Trivedi. 2003. "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.

Nghiem, Son, and Nicholas Graves. 2019. "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.

Reschovsky, James D, Peter Kemper, and Ha Tu. 2000. "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.

Ritter, Christian, and Martin A Tanner. 1992. "Facilitating the Gibbs Sampler: the Gibbs Stopper and the Griddy-Gibbs Sampler." *Journal of the American Statistical Association*, 87(419): 861–868.

Sapelli, Claudio, and Bernardita Vial. 2003. "Self-Selection and Moral Hazard in Chilean Health Insurance." *Journal of Health Economics*, 22(3): 459–476.

Vossmeyer, Angela. 2016. "Sample Selection and Treatment Effect Estimation of Lender of Last Resort Policies." *Journal of Business & Economic Statistics*, 34(2): 197–212.