Institute for Economic Studies, Keio University

Keio-IES Discussion Paper Series

Hierarchical Bayesian Hedonic Regression Analysis of Japanese Rice Wine: Price is Right?

Wakuo Saito、Teruo Nakatsuma

16 September, 2021
DP2021-019
https://ies.keio.ac.jp/en/publications/14618/

Institute for Economic Studies, Keio University
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
ies-office@adst.keio.ac.jp
16 September, 2021

Hierarchical Bayesian Hedonic Regression Analysis of Japanese Rice Wine: Price is Right?

Wakuo Saito、Teruo Nakatsuma

Keio-IES DP2021-019
16 September, 2021
JEL Classification: C13; L11; Q12
Keywords: sake; rice breed; hierarchical Bayesian modeling; hedonic pricing model; Markov chain Monte Carlo

Abstract

This research formulates a hedonic pricing model for Japanese rice wine, sake, via hierarchical Bayesian modeling, estimating it with a Markov chain Monte Carlo (MCMC) method. The data used in the estimation are obtained from Rakuten, the largest online shopping site in Japan. Flavor indicators, premium categories, rice breeds, and regional dummy variables are used as pricing factors. The Bayesian estimation of the model employs an ancillarity-sufficiency interweaving strategy to improve the sampling efficiency of MCMC. The estimation results indicate that Japanese consumers value sweeter sake more and the price reflects the cost of pre-processing rice only for the most luxurious category. No distinctive differences are identified among rice breeds or regions in the hedonic pricing model.

Wakuo Saito
Graduate School of Economics, Keio University
2-15-45, Mita, Minato-ku, Tokyo
goldcrystal1@keio.jp

Teruo Nakatsuma
Faculty of Economics, Keio University
2-15-45, Mita, Minato-ku, Tokyo
nakatsuma@econ.keio.ac.jp

Hierarchical Bayesian Hedonic Regression Analysis of Japanese Rice Wine: Price is Right?

Wakuo Saito*

Graduate School of Economics, Keio University

Teruo Nakatsuma

Faculty of Economics, Keio University

Abstract

This research formulates a hedonic pricing model for Japanese rice wine, *sake*, via hierarchical Bayesian modeling, estimating it with a Markov chain Monte Carlo (MCMC) method. The data used in the estimation are obtained from Rakuten, the largest online shopping site in Japan. Flavor indicators, premium categories, rice breeds, and regional dummy variables are used as pricing factors. The Bayesian estimation of the model employs an ancillarity-sufficiency interweaving strategy to improve the sampling efficiency of MCMC. The estimation results indicate that Japanese consumers value sweeter *sake* more and the price reflects the cost of pre-processing rice only for the most luxurious category. No distinctive differences are identified among rice breeds or regions in the hedonic pricing model.

Key words: *sake*, rice breed, hedonic pricing model, hierarchical Bayesian modeling, Markov chain Monte Carlo.

*Corresponding author: e-mail: goldcrystal1@keio.jp

1 Introduction

The worldwide spread of the novel coronavirus COVID-19 since early 2020 inflicted severe damage on the alcoholic beverage industry. Demand for alcoholic beverages such as beer and wine sharply declined across the globe due to forced closures of bars and restaurants and prohibitions on indoor dining. Japanese rice wine, *sake*, is no exception. Since the first case of COVID-19 was reported in January 2020, to prevent the spread of coronavirus nationwide, the Japanese government repeatedly declared a state of emergency, shutting down bars and restaurants or ceasing the sale of alcoholic beverages. Because bars and restaurants are major buyers of *sake*, the Japanese *sake* breweries suffered the loss of a great portion of usual sales revenue. This decline in demand for *sake* was further worsened by record-high bar and restaurant closures. According to a survey conducted by the Ministry of Agriculture, Forestry and Fisheries (MAFF), this resulted in domestic shipments of *sake* falling by 11% in 2020 from the previous year.

Under this unprecedented adverse business environment, Japanese *sake* breweries are struggling to identify alternative channels for *sake* sale. One promising alternative is e-commerce. Due to state-imposed restrictions on outside activities[1], the frequency and the volume of the online purchases of food and other necessities dramatically increased. For instance, Rakuten, Japan's largest e-commerce conglomerate, experienced solid growth in their sales revenue from the first to the fourth quarter of 2020 and it is on track to post its highest operating profit in the first quarter of 2021. Given the fact that consumers have preferred to purchase goods and have them delivered rather than leaving their homes and risking possible COVID-19infection, Japanese *sake* breweries may need to establish new online sales channels and supply more products for home consumption to compensate for the loss of bar and restaurant sales.

Although the shift to the e-commerce market seems to be a plausible strategy, its successful execution is a different matter. The Japanese *sake* industry mainly consists of family-owned small and medium-sized enterprises and their decision making continues to

[1]Unlike other countries, the Japanese government did not impose a strict lockdown upon the population. Nonetheless, the Japanese people were encouraged to limit non-essential outside activities and stay in their homes during the COVID-19 pandemic.

be based on experience and intuition. Most managers have limited expertise regarding marketing strategy in general and proper product pricing in particular. Furthermore, most *sake* breweries sell majority of their product through wholesalers and have insufficient experience in direct sale. Simply put, managers tend to follow the practices established by their parents and maintain the same old long-term relationships with wholesalers for decades. Given the prevalent old-fashioned management style in the Japanese *sake* industry, data-driven pricing of *sake* is inconceivable.

This research endeavors to assist managers of *sake* breweries who venture into the e-commerce market, proposing a hedonic pricing approach for *sake*. The application of the hedonic pricing approach in the liquor market is not new, especially for the wine market. For instance, seminal studies, such as Nerlove (1995) and Combris et al. (1997), concluded that the rank of vintage, wine color, and the amount of sugar in wine had significant effects on the price of wine. Costanigro (2006), Galizzi (2007), Corsi and Strom (2013), and Brentari et al. (2014), among others, determined that grape production areas had a significant impact on price. Corsi and Storm (2013) also found that consumers tended to purchase wine at a higher price if it was manufactured using organic farming. Notably, Galati et al. (2017) conducted a hedonic analysis of the Japanese wine market. Nonetheless, in comparison to the wine market, virtually no hedonic analyses of the Japanese *sake* market have been conducted as far as the authors knows.

As noted in Section 2, the *sake* brewing process differs from that of beer or wine, which necessitates the identification of key factors that specifically determine the price of *sake*. Possible candidates for explanatory variables in the hedonic pricing model include

- product categories: super premium, premium, or regular;
- breeds of rice used in the *sake* brewing process;
- brewing experts supervising the brewing process; and
- specific flavor characteristics of *sake*.

The influence of each variable on the quality of *sake* is examined in Section 2.

In the literature regarding the hedonic pricing approach, pricing models are often supposed to be a linear regression of the log price. This convention is followed for this investigation of *sake* pricing. Although it is tempting to estimate such a linear regression with the ordinary least squares (OLS) method, in our experience, OLS elicited unstable estimation results. This is mainly because the hedonic pricing model used in this study contains numerous dummy variables related to product categories and producing prefectures of *sake*. Therefore, rather than using OLS, the hedonic pricing model is constructed via hierarchical Bayesian modeling to provide more stable estimates of the coefficients in the model. Because hierarchical Bayesian modeling renders the hedonic pricing model rather complicated, the Markov chain Monte Carlo (MCMC) method and the ancillarity-sufficiency interweaving strategy (ASIS) are used for estimation.

The organization of this paper is as follows. Section 2 describes the basic information regarding *sake* for those who are unfamiliar with this traditional Japanese liquor, outlining *sake*'s unique flavor characteristics. The data set used to estimate the hedonic pricing model is also introduced. Section 3 presents the hierarchical Bayesian modeling of *sake* prices and outlines the Bayesian MCMC estimation procedure, with more details regarding this approach provided in appendices. Section 4 delineates the hypotheses tested and interprets the estimation results. Finally, concluding remarks are shared in Section 5.

2 Flavor Determinants and Other Factors for *Sake* Pricing

Prior to introducing the candidates for explanatory variables in the hedonic pricing regression of *sake*, we will now describe the key aspects of the brewing process and the determinants of the flavor. *Sake* is a traditional Japanese liquor brewed from rice. It is slightly yellow-colored, which is similar to white wine, and contains 13 to 16% alcohol. *Sake* is made from rice, *koji*, yeast, and water. Some breweries add brewed alcohol to *sake* as a post-production flavor enhancement to lower production costs. This study does not include this type of *sake* because such *sake* is a mass-produced cheap liquor that may not be suitable for analyzing the relationship between price and the quality. As such, this research focuses on *sake* without the post-production addition of alcohol, which is known as *junmai*, which means “pure rice” in Japanese.

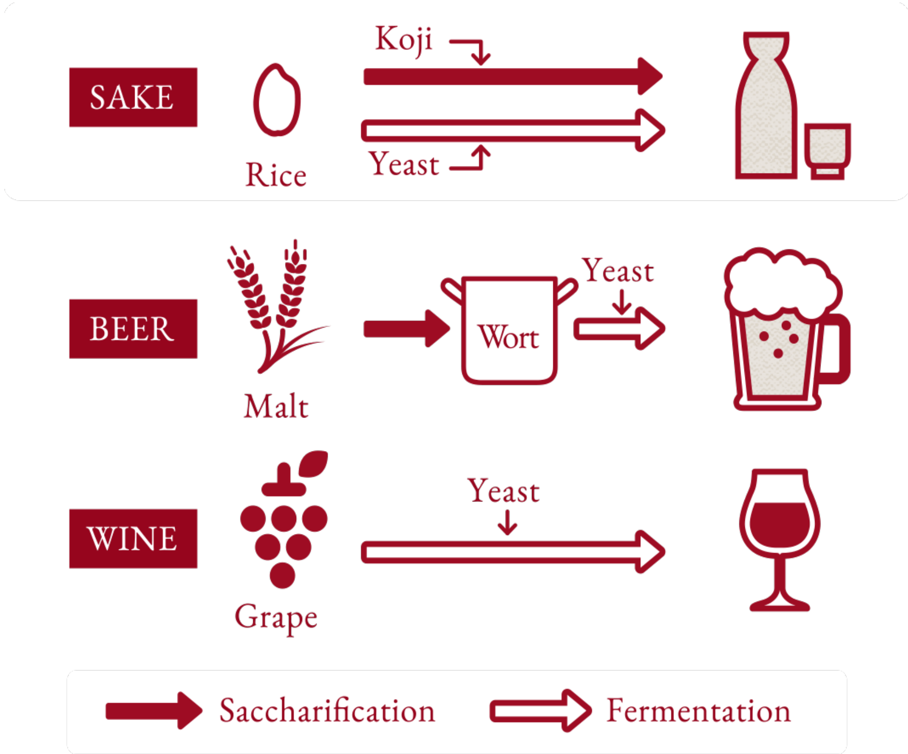

Figure 1: Brewing Processes of Each Form of Liquor[2]

One of the key materials, *koji*, is a kind of mold that decomposes rice starch into sugar. This process of sugar creation is called saccharification. The addition of yeast produces alcohol from the sugar created by *koji*, in a process called fermentation. As shown in Figure 1, saccharification and fermentation in the *sake* brewing process proceed in parallel. This parallel fermentation generates a unique flavor known as *umami*, which is created by a rich amount of amino acids. In contrast, in the beer brewing process, malt is first saccharified into wort, and this wort is fermented into alcohol with the help of yeast. This serial fermentation is illustrated in the center of Figure 1. Unlike *sake* or beer, the wine brewing process does not require saccharification because the yeast can use the grape sugar for fermentation, as shown in the bottom of Figure 1.

Figure 2 summarizes the four flavor components of *sake*, beer, and wine. In each panel of Figure 2, the top bar is for wine, the middle bar is for beer ,and the bottom bar is for *sake*. The upper-left panel shows the amount of alcohol by percentage, which

[2]The authors are grateful to the National Research Institute of Brewing for kindly providing Figure 1 and permission for its reprint.

is often referred to as alcohol by volume (ABV) in the Japanese *sake* industry, showing that *sake* has the highest ABV. The upper-right panel with the title "Extract" presents a bar chart regarding the amount of sugar by percentage. There is no distinctive difference among three liquors, other than variation being highest for wine and the average level being lower for beer. The bottom-left panel shows that *sake* contains more amino acids than the other types of liquor, but the acidity of *sake* is lower than that of wine, as shown in the bottom-right panel. Given these observations, the following indicators were chosen as explanatory variables in the hedonic pricing regression[3]:

- ABV
- *sake* meter value (SMV)
- acidity

SMV is related to the amount of sugar in *sake* and takes either a positive or negative value. A higher SMV indicates a lower amount of sugar; thus *sake* with a higher (lower) SMV tastes drier (sweeter).

At the beginning of the brewing process of *sake*, grains of rice are threshed and polished so that only the inner part of a rice grain is used for saccharification and fermentation. This is because the inner part contains more amount of starch than the peripheral part. Polishing rice grains further makes *sake* taste smoother. The downside of this polishing process is that it increases production cost by discarding a substantial portion of rice that could otherwise be used. Thus, the portion of rice that is polished is a key factor that determines both the flavor of *sake* and its production cost. This variable is measured through a polishing rice ratio (PRR); for instance, 50% PRR indicates that the outer half portion of the rice grain is discarded.

Related to PRR, the MAFF of Japan prescribes three categories of *sake*, including *jumnai*, *junmai ginjo*, and *junmai dai ginjo*. As previously noted, *junmai* is made from rice without the addition of post-production alcohol. There is no specific requirement on

[3]Although it is preferable to include the amount of amino acids in the hedonic pricing regression, this consideration was excluded from the study due to limitations in data availability.

[4]The authors appreciate the National Research Institute of Brewing for kindly supplying Figure 2 and allowing its reprint.

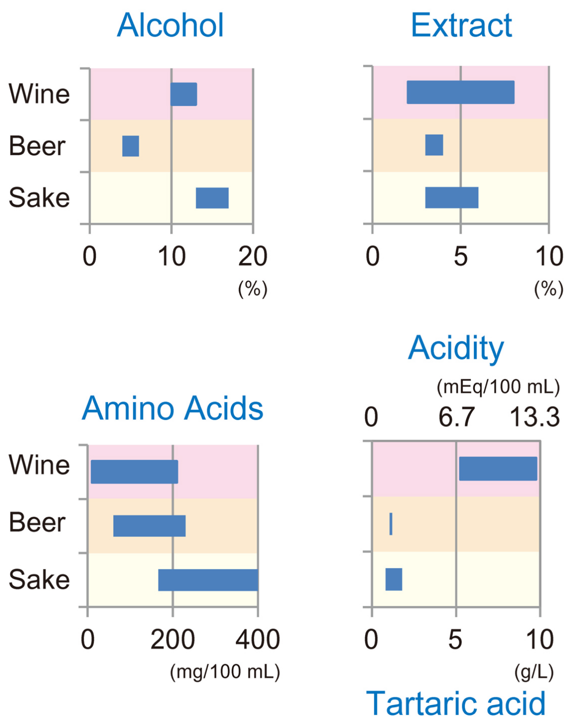

Figure 2: Liquor Flavor Comparison[4]

PRR to be classified as *junmai*; thus, it is regarded as a regular type of *sake* made purely from rice. The next category *junmai ginjo* is destined for the premium *sake* market. *Ginjo* literally means "premium brewing" in Japanese. To be categorized as *junmai ginjo*, the *sake* must be *junmai*, i.e., without the addition of post-production alcohol, and the PRR must be no more than 60%. The last *junmai dai ginjo* designation represents the super-premium category. *Dai* means "great" in Japanese. To be classified as *junmai dai ginjo*, the PRR must be no more than 50%. For convenience, the following abbreviations will be used for these categories:

- *junmai* (JM)
- *junmai ginjo* (JG)
- *junmai dai ginjo* (DG)

Given the above categorization, DG is expected to be the most expensive luxury *sake*, followed by JG, and JM to be the cheapest. To analyze how this categorization con-

Table 1: Classes of Pure Rice *Sake*

Class	*Junmai Dai Ginjo* (DG)	*Junmai Ginjo* (JG)	*Junmai* (JM)
Maximum PRR	50%	60%	None

Source: Akiyama (1994) pp. 4–7

tributes the price differentiation among JM, JG, and DG, dummy variables are used in the hedonic pricing regression to represent these categories. Cross-product terms between each category dummy and PRR are also added to the regression to assess differences in sensitivity to PRR among categories.

Unlike mass-produced cheap *sake*, the *ginjo*-type premium *sake* (DG and JG) is painstakingly crafted by a *toji* (brew master) who supervises the entire brewing process. Since most *sake* breweries are family-owned small businesses, they cannot afford to hire their own brew masters, so they outsource the task to freelance brew masters who have traditionally formed independent guilds. According to the Japan Sake and Shochu Makers Association (2021), there are 19 such guilds in Japan. In the past, knowledge and skills in *sake* brewing were a tightly held secret from the public; thus, making it necessary for breweries to hire brew masters from one of the existing guilds each season. This tradition remains today, even after an AI-monitored brewing system and other innovations have made it possible to produce *ginjo*-type premium *sake* without the help of traditional brew masters. Therefore, the quality of *sake* can differ region-to-region because it is subject to the skills and the preferences of the brew masters who belong to different regional guilds. Furthermore, some argue that regional climate differences may also affect the quality of *sake*, though with the introduction of fully automated temperature and humidity control in the brewing process, this may no longer be the case. Regional dummy variables have been introduced into the hedonic pricing regression to investigate these parameters.

Rice breeds are included as the final factor in the hedonic pricing regression of *sake*. Of course, rice is the most important material of *sake*. Although some breeds of cooking rice have been used for *sake* brewing, *ginjo*-type premium *sake* is almost exclusively made from a *sake*-specific breed of rice, which is called *sakamai* (*sake* rice) in Japanese.

One of the most commonly used *sakamai* is *Yamadanishiki*, which is mainly grown in the western Japan, whereas rice farmers in the eastern Japan mainly grow *Gohyakumangoku*. *Omachi* and *Miyamanishiki* are also popular breeds, though they are cropped in smaller quantities than *Yamadanishiki* or *Gohyakumangoku*. *Omachi* is primarily cultivated in the western regions, while *Miyamanishiki* is mostly grown in the eastern regions. Moreover, some locally grown breeds of *sakamai* are also used by local breweries in *sake* brewing, but it is difficult for non-local breweries to purchase such breeds. Dummy variables for the four major breeds of *sakamai* are included as explanatory variables in the hedonic pricing model:

- *Yamadanishiki* (YM)
- *Gohyakumanngoku* (GH)
- *Omachi* (OM)
- *Miyamanishiki* (MY)

When all four dummy variables equal zero, this means that the corresponding *sake* is produced from a locally grown breed of *sakamai*.

In summary, the following explanatory variables are included in the hedonic pricing regression model of *sake* in this study.

Flavor indicators:

- PRR
- ABV
- *sake* meter value (SMV)
- acidity

Premium categories:

- *junmai ginjo* (JG) dummy
- *junmai dai ginjo* (DG) dummy
- JG dummy × PRR

- DG dummy × PRR

Rice breeds:

- *Yamadanishiki* (YM) dummy
- *Gohyakumanngoku* (GH) dummy
- *Omachi* (OM) dummy
- *Miyamanishiki* (MY) dummy

Regional effects:

- prefecture dummies for the following 29 prefectures:

Hokkaido	*Aomori*	*Miyagi*	*Akita*	*Yamagata*	*Fukushima*
Ibaragi	*Tochigi*	*Gunma*	*Saitama*	*Chiba*	*Niigata*
Ishikawa	*Fukui*	*Nagano*	*Gifu*	*Shizuoka*	*Aichi*
Mie	*Shiga*	*Osaka*	*Hyogo*	*Nara*	*Wakayama*
Shimane	*Okayama*	*Hiroshima*	*Yamaguchi*	*Kochi*	

The data for the above variables, along with *sake* prices, was obtained on August 6, 2021, using an API provided by Rakuten and retrieving 403 observations. Note that the data set used here only reflects only information on how online retailers who operate on Rakuten's online shopping site set the prices of their products. This is a notable limitation. The descriptive statistics of *sake* prices, PRR, ABV, SMV, and acidity are summarized in Table 2. As expected, DG is the most expensive and has the lowest PRR and acidity; whereas, JM is the cheapest and has the highest PPR and acidity. JG is somewhere in between. SMV is lower for DG and JG than JM. These observations suggest that premium *sake* such as DG and JG tends to taste sweeter than less expensive JM. As for ABV, no significant differences are noted among the three categories.

Table 2: Descriptive Statistics

	Price	PRR	ABV	SMV	Acidity	Price	PRR	ABV	SMV	Acidity
			Total						**DG**	
Mean	2394	0.53	15.8	2.36	1.56	3860	0.43	15.9	1.25	1.44
Â SD	1924	0.09	0.9	5.12	0.31	2788	0.07	0.77	4.74	0.23
Â Max	22000	0.8	19	27	3.6	22000	0.5	18	13	2.5
Â Min	1034	0.18	11	-36	1	1365	0.18	14	-36	1
			JG						**JM**	
Mean	1782	0.55	15.9	1.98	1.6	1500	0.61	15.5	4.18	1.66
Â SD	283	0.03	0.84	5.1	0.32	235	0.05	1.05	5.14	0.35
Â Max	2992	0.6	18	20	3.6	2536	0.8	19	27	3.5
Â Min	1078	0.45	13	-21	1	1034	0.5	11	-20	1.1

3 Hierarchical Bayesian Modeling of the Hedonic Pricing Regression

This section will first introduce the hierarchical Bayesian modeling of the hedonic pricing regression of *sake* developed for this research. Suppose y_i is the log price of *sake* brand $i \in \{1, \dots, N\}$ and a dummy variable is defined as $d_{\star j}^{(i)}$, $\star \in \{R, B\}$,

$$d_{\star j}^{(i)} = \begin{cases} 1, & \text{if } \{\text{(R)egion, (B)reed}\} \text{ of } \textit{sake} \text{ brand } i \text{ is } j; \\ 0, & \text{otherwise,} \end{cases}$$

where "Region" refers to the prefecture where sake brand i is produced and "Breed" represents the rice breed from which the *sake* brand is made. As for the regional dummy variable d_{Rj}^i, all dummy variables for 29 prefectures are included. For this reason, the constant term from the regression is excluded to avoid multicollinearity. As for the breed dummy variable d_{Bj}^i, any local rice breeds are treated as the base breed; that is, $d_{\star j}^i = 0$ for all j if *sake* brand i is made from a local rice breed. Further, suppose $\{x_{ki}\}_{k=1}^K$ are explanatory variables, including flavor indicators (PRR, ABV, SMV, and acidity), dummy variables for premium categories (DG dummy and JG dummy), and the cross-product terms (DG dummy×PRR and JG dummy×PRR). The hedonic pricing regression model is formulated as

$$y_i = \sum_{j=1}^{N_R} d_{Rj}^{(i)} \alpha_{Rj} + \sum_{j=1}^{N_B} d_{Bj}^{(i)} \alpha_{Bj} + \sum_{k=1}^{K} x_{ik} \beta_k + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2), \tag{1}$$

where $N_R = 29$ (the number of prefectures analyzed), $N_B = 4$ (the number of rice breeds), and $K = 8$ (the number of other explanatory variables) in the study. By introducing the following notations:

$$\boldsymbol{D}_\star = \begin{bmatrix} d_{\star 1}^{(1)} & \cdots & d_{\star N_\star}^{(1)} \\ \vdots & \ddots & \vdots \\ d_{\star 1}^{(N)} & \cdots & d_{\star N_\star}^{(N)} \end{bmatrix}, \quad \boldsymbol{\alpha}_\star = \begin{bmatrix} \alpha_{\star 1} \\ \vdots \\ \alpha_{\star N_\star} \end{bmatrix}, \quad \star \in \{R, B\},$$

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}, \quad \boldsymbol{X} = \begin{bmatrix} x_{11} & \cdots & x_{1K} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{NK} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_K \end{bmatrix}, \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_N \end{bmatrix},$$

the regression model (1) is rewritten as

$$\boldsymbol{y} = \boldsymbol{D}_R \boldsymbol{\alpha}_R + \boldsymbol{D}_B \boldsymbol{\alpha}_B + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}\left(\boldsymbol{0}_N, \sigma_\epsilon^2 \boldsymbol{I}_N\right), \tag{2}$$

where $\boldsymbol{0}_N$ is the $N \times 1$ zero vector and $\boldsymbol{I}_N$ is the $N \times N$ identity matrix. Similar expressions will be used for zero vectors and identity matrices with different shapes. Furthermore, by defining

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{D}_R & \boldsymbol{D}_B & \boldsymbol{X} \end{bmatrix}, \quad \boldsymbol{\delta} = \begin{bmatrix} \boldsymbol{\alpha}_R \\ \boldsymbol{\alpha}_B \\ \boldsymbol{\beta} \end{bmatrix},$$

we have

$$\boldsymbol{y} = \boldsymbol{Z}\boldsymbol{\delta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}\left(\boldsymbol{0}_N, \sigma_\epsilon^2 \boldsymbol{I}_{NT}\right). \tag{3}$$

Following previous studies, we initially tried to estimate (3) using the OLS method, but in our experience, this conventional method did not elicit stable estimation results.

Instead, this study estimates (3) via hierarchical Bayesian modeling. For the above model (3), the likelihood of unknown parameters $(\boldsymbol{\delta}, \sigma_\epsilon)$ is given by

$$p(\boldsymbol{y}|\boldsymbol{Z}, \boldsymbol{\delta}, \sigma_\epsilon) \propto (\sigma_\epsilon^2)^{-\frac{N}{2}} \exp\left[-\frac{1}{2\sigma_\epsilon^2}(\boldsymbol{y} - \boldsymbol{Z\delta})^\intercal(\boldsymbol{y} - \boldsymbol{Z\delta})\right] \tag{4}$$

$$\propto (\sigma_\epsilon^2)^{-\frac{N}{2}} \exp\left[-\frac{\sum_{i=1}^N e_i^2}{2\sigma_\epsilon^2}\right], \tag{5}$$

where $e_i = y_i - \sum_{j=1}^{N_R} d_{Rj}^{(i)}\alpha_{Rj} - \sum_{j=1}^{N_B} d_{Bj}^{(i)}\alpha_{Bj} - \sum_{k=1}^{K} x_{ik}\beta_k$. Two different forms of the likelihood, (4) and (5), will be used to derive the conditional posterior distribution of each parameter.

The prior distribution of $\boldsymbol{\delta}$ and σ_ϵ is assumed to be:

$$\boldsymbol{\delta} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_R \mathbf{1}_{N_R} \\ \mu_B \mathbf{1}_{N_B} \\ \boldsymbol{\mu}_\beta \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_R^2 \boldsymbol{I}_{N_R} & & \\ & \sigma_B^2 \boldsymbol{I}_{N_B} & \\ & & \boldsymbol{\Sigma}_\beta \end{bmatrix}, \tag{6}$$

$$\sigma_\epsilon \sim \mathcal{C}^+(0, s_\epsilon), \tag{7}$$

where $\mathcal{C}^+(0, s_\epsilon)$ stands for the half-Cauchy distribution:

$$p(\sigma_\epsilon|s_\epsilon) = \frac{2s_\epsilon}{\pi(\sigma_\epsilon^2 + s_\epsilon^2)}, \quad \sigma_\epsilon > 0, \quad s_\epsilon > 0, \tag{8}$$

and s_ϵ takes a preset value as a hyper-parameter. Note that the prior distribution (6) is equivalent to

$$\alpha_{\star i} \sim \mathcal{N}(\mu_\star, \sigma_\star^2), \quad i \in \{1, \dots, N_\star\}, \quad \star \in \{R, B\}, \tag{9}$$

$$\boldsymbol{\beta} \sim \mathcal{N}(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta), \quad j \in \{1, \dots, K\}. \tag{10}$$

We further assume the prior distribution of $\mu_\star$ and $\sigma_\star$ in (6) are

$$\mu_\star \sim \mathcal{N}(\varphi_\star, \tau_\star^2), \quad \sigma_\star \sim \mathcal{C}^+(0, s_\star), \quad \star \in \{R, B\}, \tag{11}$$

where ($\varphi_\star, \tau_\star$, and $s_\star$) are also hyper-parameters, which are fixed at preset values. Gelman (2006) suggests that the half-Cauchy distribution (8) is more suitable as the prior distribution for the variance parameter in a hierarchical model such as $\sigma_\star$ in (11). Finally, prior distributions (6), (7), (9) – (11) are summarized into the joint prior distribution of

$\boldsymbol{\theta} = (\boldsymbol{\delta}, \mu_R, \sigma_R, \mu_B, \sigma_B, \sigma_\epsilon)$:

$$
\begin{aligned}
p(\boldsymbol{\theta}) = p(&\boldsymbol{\delta}|\mu_R, \sigma_R, \mu_B, \sigma_B, \boldsymbol{\mu}_\beta \boldsymbol{\Sigma}_\beta) \\
&\times p(\mu_R|\varphi_R, \tau_R) p(\sigma_R|s_R) p(\mu_B|\varphi_B, \tau_B) p(\sigma_B|s_B) p(\sigma_\epsilon|s_\epsilon).
\end{aligned}
\tag{12}
$$

For brevity in mathematical expressions, dependency on the hyper-parameters was ignored in the prior distribution (12) as $p(\boldsymbol{\theta})$. By applying Bayes' theorem to the likelihood (4) and prior distribution (12), the posterior distribution of $\boldsymbol{\theta}$ is:

$$
p(\boldsymbol{\theta}|\mathcal{D}) \propto p(\boldsymbol{y}|\boldsymbol{Z}, \boldsymbol{\delta}, \sigma_\epsilon) p(\boldsymbol{\theta}), \quad \mathcal{D} = (\boldsymbol{y}, \boldsymbol{Z}). \tag{13}
$$

Unfortunately, the posterior distribution (13) and the posterior statistics of $\boldsymbol{\theta}$ cannot be analytically evaluated as moments and intervals. Instead, they are evaluated using the MCMC method. The results indicate that the conditional posterior distributions of all parameters are readily available; thus, a Gibbs sampler algorithm method can be adopted to generate pseudo-random numbers of $\boldsymbol{\theta}$ from the posterior distribution (13) for applying the Monte Carlo integration to evaluate the posterior statistics, including the posterior mean, the posterior standard deviation, and the interval estimation. In the Gibbs sampler, ($\boldsymbol{\delta}, \mu_R, \sigma_R, \mu_B, \sigma_B$, and σ_ϵ) are generated one by one in Steps 1 – 6.

Gibbs sampler for the hierarchical Bayesian regression model

Step 1. Draw $\boldsymbol{\delta}$ from the conditional posterior distribution $p(\boldsymbol{\delta}|\mathcal{D}, \boldsymbol{\theta}_{-\boldsymbol{\delta}})$.

Step 2. Draw μ_R from the conditional posterior distribution $p(\mu_R|\mathcal{D}, \boldsymbol{\theta}_{-\mu_R})$.

Step 3. Draw σ_R from the conditional posterior distribution $p(\sigma_R|\mathcal{D}, \boldsymbol{\theta}_{-\sigma_R})$.

Step 4. Draw μ_B from the conditional posterior distribution $p(\mu_B|\mathcal{D}, \boldsymbol{\theta}_{-\mu_B})$.

Step 5. Draw σ_B from the conditional posterior distribution $p(\sigma_B|\mathcal{D}, \boldsymbol{\theta}_{-\sigma_B})$.

Step 6. Draw σ_ϵ from the conditional posterior distribution $p(\sigma_\epsilon|\mathcal{D}, \boldsymbol{\theta}_{-\sigma_\epsilon})$.

Note that $\boldsymbol{\theta}_{-x}$ indicates that a parameter x is excluded from $\boldsymbol{\theta}$. Each step generates a new value of the parameter from the conditional posterior distribution, replacing the current value with the new one before moving on to the next step. The loop of Steps 1-6 is started from an arbitrary initial point of $\boldsymbol{\theta}$ and repeated until the generated sample paths of the parameters are stabilized. This initial sampling is known as “burn-in” in the literature. In our experience, the plain vanilla Gibbs sampler tends to generate highly correlated unstable sample paths, which may be caused by the fact that the hedonic pricing regression (3) includes many dummy variables. Therefore, to improve the efficiency of random number generation in the Gibbs sampler, the ASIS proposed by Yu and Meng (2011) is applied so that the sample paths of the parameters generated will be stabilized faster. See the Appendix for more information on the derivation of each conditional posterior parameters and the ASIS method.

4 Results

To establish hypotheses for statistical inference on the relationship between the price of *sake* and the potential candidates for determinants of the quality presented in Section 2, we interviewed Professor Tsutomu Fujii[5], who is currently affiliated with Faculty of Food and Agricultural Sciences, Fukushima University, and was the supervisor of the Department of Quality and Evaluation Research Division in the National Research Institute of Brewing. In his former career, he evaluated the quality of various kinds of *sake* as a judge for the Annual Japan Sake Awards, which is the most traditional and prestigious *sake* competition. Based on his knowledge and experience, Professor Fujii suggested the following “conventional wisdom” in the *sake* industry related to the signs of the coefficients in the hedonic pricing regression:

Flavor indicators:

- H1 The coefficient for PRR will be negative because lowering PRR costs more.
- H2 The coefficient for ABV will be positive because a higher ABV is an essential factor for the fragrance of DG and JG.
- H3 The coefficient for SMV will be negative because lower SMV leads to higher quality for *junmai*.
- H4 The coefficient for the acidity will be negative. If the acidity is higher than 1.7, such *sake* is no longer classified in DG or JG.

[5]Details on his academic achievements are available at https://researchmap.jp/read0005781

Premium categories:

H5 The coefficient for JG dummy should be positive because of the PRR cap (it must be no more than 60%), as noted in Section 2.

H6 The coefficient for DG dummy should also be positive for the same reason as H5.

H7 For both the JG dummy × PRR and the DG dummy × PRR, the coefficient will be negative for the same reason as H5 and H6.

Rice breeds:

H8 The *Yamadanishiki* (YM) dummy should be positive and have the highest impact because YM is the most suitable *sakamai* for brewing DG and JG.

H9 The *Gohyakumanngoku* (GH) dummy will not have a high impact.

H10 The *Omachi* (OM) dummy will have a high positive impact next to YM.

H11 The *Miyamanishiki* (MY) dummy will not have a significant impact.

Regional effects:

H12 There will be no clear difference among prefectures regarding regional effects because contemporary brewing technologies are almost universally used throughout Japan, as noted in Section 2.

In the Gibbs sampler, the hyper-parameters in the prior distributions (7), (9) – (11) were set as

$$\boldsymbol{\mu}_\beta = \mathbf{0}_K, \quad \boldsymbol{\Sigma}_\beta = 100\boldsymbol{I}_K, \quad s_\epsilon = 1,$$
$$\varphi_\star = 0, \quad \tau_\star^2 = 100, \quad s_\star = 1, \quad \star \in \{R, B\}.$$

The number of the initial burn-in iterations for the Gibbs sampler was 5,000, and then we generated 50,000 sets of parameters from the posterior distribution (13).

Table 3 presents the estimation results via hierarchical Bayesian modeling. This table includes the names of variables, point estimates (posterior mean) of the coefficients, the posterior standard deviations of the coefficients as "SD", and the 90% intervals[6] as "90%".

[6]We use the highest posterior density interval for interval estimation.

Table 3: Estimation Results

Variables	Coefficients	SD	90% Interval	Variables	Coefficients	SD	90% Interval
PRR	-0.355	0.502	[-1.195,0.457]	*Saitama*	7.206	0.417	[6.533,7.905]
ABV	0.021	0.016	[-0.007,0.047]	*Chiba*	7.230	0.411	[6.553,7.905]
SMV	-0.006	0.003	[-0.011,-0.001]	*Niigata*	7.284	0.407	[6.607,7.944]
Acidity	0.011	0.049	[-0.068,0.091]	*Ishikawa*	7.287	0.413	[6.601,7.954]
JG	0.180	0.465	[-0.568,0.962]	*Fukui*	7.183	0.412	[6.496,7.849]
DG	2.484	0.346	[1.906,3.047]	*Nagano*	7.160	0.411	[6.485,7.837]
JG×PRR	-0.097	0.808	[-1.443,1.212]	*Gifu*	7.171	0.406	[6.506,7.840]
DG×PRR	-4.128	0.614	[-5.155,-3.136]	*Shizuoka*	7.248	0.410	[6.561,7.910]
YM	0.024	0.033	[-0.025,0.082]	*Aichi*	7.178	0.411	[6.492,7.843]
GH	-0.014	0.037	[-0.077,0.044]	*Mie*	7.216	0.411	[6.534,7.885]
OM	0.022	0.044	[-0.046,0.098]	*Shiga*	7.210	0.411	[6.536,7.888]
MY	-0.039	0.044	[-0.111,0.027]	*Osaka*	7.235	0.414	[6.565,7.925]
Hokkaido	7.261	0.402	[6.594,7.916]	*Hyogo*	7.259	0.411	[6.594,7.948]
Aomori	7.216	0.411	[6.560,7.911]	*Nara*	7.223	0.413	[6.548,7.906]
Miyagi	7.204	0.407	[6.526,7.864]	*Wakayama*	7.200	0.408	[6.514,7.859]
Akita	7.203	0.410	[6.534,7.883]	*Shimane*	7.214	0.413	[6.531,7.891]
Yamagata	7.097	0.407	[6.434,7.774]	*Okayama*	7.172	0.413	[6.498,7.857]
Fukushima	7.183	0.408	[6.511,7.855]	*Hiroshima*	7.223	0.412	[6.560,7.913]
Ibaragi	7.205	0.409	[6.529,7.878]	*Yamaguchi*	7.142	0.412	[6.459,7.814]
Tochigi	7.239	0.41	[6.562,7.907]	*Kochi*	7.230	0.407	[6.553,7.890]
Gunma	7.156	0.413	[6.488,7.845]				

As the confidence interval in OLS estimation, the sign of the coefficient is inferred to be inconclusive if the corresponding 90% interval includes zero. Conversely, if the entire 90% interval is on the positive (negative) region, we conclude that the corresponding coefficient is positive (negative).

First, the hypotheses regarding flavor indicators (H1 – H4) are tested. The point estimate of PRR is negative, while that of ABV is positive. Although these estimates are consistent with H1 and H2, their signs are inconclusive because the 90% interval includes zero for both cases. The coefficient for SMV is conclusively negative, supporting H3. The coefficient for acidity is negative, but it is inconclusive because the 90% interval includes zero, which means that H4 is not supported. These results imply that lower SMV (sweeter *sake*) is more valued in the online market but other flavor indicators have negligible impact on price.

Next, H5 – H7, which are related to the influence of premium categories on the price, are examined. In Table 3, the sign of the JG dummy coefficient is ambiguous, but that of DG dummy is positive and substantial, so H6 is supported, but H5 is not. As for H7, the sign of the coefficient of the cross-term JG×PRR is inconclusive but that of DG×PRR is conclusively negative. Therefore, as "super premium" *sake*, DG seems to have a distinctive PRR-price profile, in which the intercept is positive (DG is more expensive than JM and JG), and the slope is negative (lower PRR leads to a higher price).

As for rice breeds, none of the four dummy variables, YM, GH, OM, and MY, elicited a conclusively positive or negative coefficient; thus, H8 and H10 are not supported, while H9 and H11 are somewhat consistent with the data.

Finally, regional effects from *Hokkaido* to *Kochi* are compared in Table 3. All estimates are positive and range from 7.0 to 7.3, but no statistically noticeable differences are found among them; hence, H12 is supported.

In summary, the estimation results in Table 3 suggest the following findings.

1. Lower SMV leads to a higher price in general, which may indicate that Japanese consumers prefer sweeter *sake*.

2. As it is categorized as DG, "super premium" *sake* has a strongly positive impact

on the price.

3. DG with lower PRR tends to be priced higher, which may reflect the cost of the polishing process in addition to flavor improvement.

4. Both rice breed and producing prefecture have negligible impact on price.

5 Conclusions

This research estimated a hedonic pricing model for Japanese rice wine, *sake*, with data obtained from Rakuten's online shopping site. Flavor indicators, premium categories, rice breed, and regional dummies were used as explanatory variables in the hedonic pricing regression as possible determinants of *sake* prices. To obtain more stable estimation results, the hedonic pricing model was constructed via hierarchical Bayesian modeling, and the model was estimated using the MCMC method. ASIS was used to enhance the efficiency of the sampling algorithm.

In the estimated hedonic pricing model, the amount of sugar, which is negatively related to SMV, had a positive impact on price; thus it can be inferred that Japanese consumers prefer sweeter *sake*. PRR has a negative impact on the price only if the *sake* is categorized as *junmai dai ginjo* (DG) "super premium" *sake*. This may imply that the costly polishing process is justified only for the most luxury category. DG was also found to be priced higher than other less luxury *sake*. Although some flavor indicators seem to influence *sake* prices, rice breeds and producing prefectures appear to have little to do with them.

COVID-19 still threatens the *sake* brewing industry in Japan. The Japanese government adheres to "lockdown" measures and vaccination requirements to suppress the spread of the virus, and as a result, bar and restaurant revenues have not yet recovered to pre-pandemic levels. We believe that a shift to concentrate on the e-commerce market is vital, and a proper pricing strategy is essential for the *sake* brewing industry. We hope that our research findings will be of some help for the industry.

References

[1] Akiyama, H. (1994) *"Nihonshu" Iwanami Shoten* (in Japanese).

[2] Brentari, E, Levaggi, R, and Zuccolotto, P. (2014) "A Hedonic Price Analysis for the Italian Wine in the Domestic Market" *Quality & Quantity: International Journal of Methodology*, Vol. 49, No.3, pp. 999-1012

[3] Combris, P, Lecocq, S, and Visse, M. (1997) "Estimation of a Hedonic Price Equation for Bordeaux Wine: Does Quality Matter?" *The Economic Journal*, Vol. 107, No.441, pp. 390-402.

[4] Constanigro, M, McCulskey, J, and Mittelhammer, R. (2007) "Segmenting the Wine Market Based on Price: Hedonic Regression When Different Prices Mean Different Products," *The Journal of Agricultural Economics*, Vol.58(3) , pp. 454-466.

[5] Corsi, A and Storm, S. (2013) "The Premium for Organic Wines: Estimating a hedonic price farm-gate price equation," The Journal of Wine economics. Vol 8, No.1, pp. 28-48.

[6] Gelman, A. (2006) *"Data Analysis Using Regression and Multilevel/Hierarchical Models"* Cambridge University Press.

[7] Gelman, A, Carlin, J, Stern, H, Dunson.D, Vehtari, and Rubin.D. (2013) *"Bayesian Data Analysis, Third Edition"*, CRC Press.

[8] Galati, A, Crescimanno, M, and Tinervia, S. (2017) "Italian Red Wine in the Japanese Market: a Hedonic Price Analysis" *Global Business and Economics Review*, Vol.19, No.6, pp. 760-770.

[9] Hoang, V, Iida, T, Matsumoto, S,Watanabe, N, and Wilson, C. (2016) "Consumer's Comparison Between Local and Imported Organic Products: a Hedonic Analysis of the Japanese Table Wine Market," *Eurasian Business Review*, Vol. 6, No,3 ,pp. 405-415.

[10] Nerlove, M. (1995) ”Hedonic Price Functions and the Measurement of Preferences; the Case of Swedish Wine Consumers,” *European Economic Review*, Vol.39, No.9, pp. 1967-1716.

[11] Steiner, B. (2004) ”Valuing Labeling Attributes with Hedonic Price Analysis: Australian Wines in the British Wine Market,” *Agribusiness*, Vol.20, No.3, pp. 287-307.

[12] Schmidt, D.F and Makalic,E. (2016) ”A Simple Sampler for the Horseshoe Estimator,” *IEEE Signal Processing Letters* Vol.23, No.1, pp.179-182.

[13] Rosen, S.M. (1974) ”Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition,” *Journal of Political Economy*, Vol.82, No.1, pp. 34-55.

[14] Yu, Y. and Meng, X.L. (2011) ”To Center or Not to Center: That is Not the Question - An Ancillarity-Sufficiency Interweaving Strategy (ASIS) for Boosting MCMC Efficiency,” *Journal of Computational and Graphical Statistics*, Vol.20, pp. 531- 570.

[15] Wand, M.P, Ormerod, J.T, Padoan, S.A, FrÃŒhwirth, R. (2011) ”Mean Field Variational Bayes for Elaborate Distributions,” *Bayesian Analysis*, Vol.6, No.4, pp. 847-900.

[16] Japan Sake and Shochu Makers Association (2021)

How is Sake Made?

https://japansake.or.jp/sake/en/basic/how-is-sake-made/. Accessed 17 August 2021

[17] Japan Sake and Shochu Makers Association (2021)

Sake Regionality.

https://japansake.or.jp/sake/en/professional/sake-regionality/. Accessed 17 August 2021

[18] National Research Institute of Brewing (2021)

Introduction to SAKE.

https://www.nrib.go.jp/sake/pdf/sake_leaflet_e.pdf. Accessed 08 September 2021

[19] Rakuten (2021) IR Results by Segment.

https://corp.rakuten.co.jp/investors/financial/segment.html. Accessed 06 August 2021

[20] The Ministry of Agriculture, Forestry and Fisheries of Japan (2021)

Shuzo Kotekimai no Jukyu Suitei Jokyo (Suikei) Reiwa 3nendo 3gatsubun.

(Supply and Demand of *sakamai* (Estimate), March 2021) (in Japanese.)

https://www.maff.go.jp/j/seisaku/_tokatu/kikaku/attach/pdf/sake_02seisan-5.pdf. Accessed 30,April 2021

Appendix: Conditional Posterior Distributions and ASIS Algorithm

In this appendix, we first derive the conditional posterior distributions of the parameters in (13) and then describe the algorithm of ASIS.

The conditional posterior distribution of $\boldsymbol{\delta}$ is derived by applying Bayes' theorem to the likelihood (4) and the prior distribution of $\boldsymbol{\delta}$ (6) as follows:

$$\begin{aligned} p(\boldsymbol{\delta}|\mathcal{D},\boldsymbol{\theta}_{-\boldsymbol{\delta}}) &\propto p(\boldsymbol{y}|\boldsymbol{Z},\boldsymbol{\delta},\sigma_\epsilon)p(\boldsymbol{\delta}|\boldsymbol{\mu},\boldsymbol{\Sigma}) \\ &\propto \exp\left[-\frac{1}{2\sigma_\epsilon^2}\left\{(\boldsymbol{y}-\boldsymbol{Z\delta})^\intercal(\boldsymbol{y}-\boldsymbol{Z\delta})\right\}\right]\times\exp\left[-\frac{1}{2}(\boldsymbol{\delta}-\boldsymbol{\mu})^\intercal\boldsymbol{\Sigma}^{-1}(\boldsymbol{\delta}-\boldsymbol{\mu})\right] \\ &= \exp\left[-\frac{1}{2}\left\{\sigma_\epsilon^{-2}(\boldsymbol{y}-\boldsymbol{Z\delta})^\intercal(\boldsymbol{y}-\boldsymbol{Z\delta})+(\boldsymbol{\delta}-\boldsymbol{\mu})^\intercal\boldsymbol{\Sigma}^{-1}(\boldsymbol{\delta}-\boldsymbol{\mu})\right\}\right]. \end{aligned} \tag{14}$$

By completing the square in (14), we have

$$\begin{aligned} &\sigma_\epsilon^{-2}(\boldsymbol{y}-\boldsymbol{Z\delta})^\intercal(\boldsymbol{y}-\boldsymbol{Z\delta})+(\boldsymbol{\delta}-\boldsymbol{\mu})^\intercal\boldsymbol{\Sigma}^{-1}(\boldsymbol{\delta}-\boldsymbol{\mu}) \\ &= \boldsymbol{\delta}^\intercal\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{Z}+\boldsymbol{\Sigma}^{-1}\right)\boldsymbol{\delta}-2\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{y}+\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right)^\intercal\boldsymbol{\delta}+\text{const} \\ &= \left(\boldsymbol{\delta}-\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{Z}+\boldsymbol{\Sigma}^{-1}\right)^{-1}\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{y}+\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right)\right)\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{Z}+\boldsymbol{\Sigma}^{-1}\right) \\ &\quad\times\left(\boldsymbol{\delta}-\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{Z}+\boldsymbol{\Sigma}^{-1}\right)^{-1}\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{y}+\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right)\right)+\text{const}, \end{aligned}$$

where "const" indicates that the term is independent of $\boldsymbol{\delta}$. Then, by dropping "const", we rearrange the conditional posterior distribution of (14) as

$$\begin{aligned} p(\boldsymbol{\delta}|\mathcal{D},\boldsymbol{\theta}_{-\boldsymbol{\delta}}) \propto \exp\Bigg[-\frac{1}{2}&\left(\boldsymbol{\delta}-\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{Z}+\boldsymbol{\Sigma}^{-1}\right)^{-1}\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{y}+\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right)\right)\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{Z}+\boldsymbol{\Sigma}^{-1}\right) \\ &\times\left(\boldsymbol{\delta}-\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{Z}+\boldsymbol{\Sigma}^{-1}\right)^{-1}\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{y}+\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right)\right)\Bigg]. \end{aligned} \tag{15}$$

(15) is rewritten as

$$\boldsymbol{\delta}|\mathcal{D},\boldsymbol{\theta}_{-\boldsymbol{\delta}} \sim \mathcal{N}\left(\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{Z}+\boldsymbol{\Sigma}^{-1}\right)^{-1}\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{y}+\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right),\left(\sigma_\epsilon^{-2}\boldsymbol{Z}^\intercal\boldsymbol{Z}+\boldsymbol{\Sigma}^{-1}\right)^{-1}\right), \tag{16}$$

which is the conditional posterior distribution $p(\boldsymbol{\delta}|\mathcal{D},\boldsymbol{\theta}_{-\boldsymbol{\delta}})$ used in the Gibbs sampler.

Next, we derive the conditional posterior distributions of $\mu_\star$, $\star\in\{R,B\}$. By applying

Bayes' theorem to (9) and (11), we have

$$\begin{aligned}
p(\mu_\star|\mathcal{D},\boldsymbol{\theta}_{-\mu_\star}) &\propto p(\boldsymbol{\alpha}_\star|\mu_\star,\sigma_\star)p(\mu_\star|\varphi_\star\tau_\star) \\
&\propto \exp\left[-\frac{\sum_{i=1}^{N_\star}(\alpha_{\star i}-\mu_\star)^2}{2\sigma_\star^2}-\frac{(\mu_\star-\varphi_\star)^2}{2\tau_\star^2}\right] \\
&\propto \exp\left[-\frac{1}{2}\left\{\left(\sigma_\star^{-2}N_\star+\tau_\star^{-2}\right)\mu_\star^2-2\left(\sigma_\star^{-2}\sum_{i=1}^{N_\star}\alpha_{\star i}+\tau_\star^{-2}\varphi_\star\right)\mu_\star\right\}\right] \\
&\propto \exp\left[-\frac{1}{2}\left(\sigma_\star^{-2}N_\star+\tau_\star^{-2}\right)\left(\mu_\star-\frac{\sigma_\star^{-2}\sum_{i=1}^{N_\star}\alpha_{\star i}+\tau_\star^{-2}\varphi_\star}{\sigma_\star^{-2}N_\star+\tau_\star^{-2}}\right)^2\right]. \qquad (17)
\end{aligned}$$

Therefore the conditional posterior distribution $p(\mu_\star|\mathcal{D},\boldsymbol{\theta}_{-\mu_\star})$ is derived as

$$\mu_\star|\mathcal{D},\boldsymbol{\theta}_{-\mu_\star} \sim \mathcal{N}\left(\frac{\sigma_\star^{-2}\sum_{i=1}^{N_\star}\alpha_{\star i}+\tau_\star^{-2}\varphi_\star}{\sigma_\star^{-2}N_\star+\tau_\star^{-2}}, \frac{1}{\sigma_\star^{-2}N_\star+\tau_\star^{-2}}\right). \qquad (18)$$

In order to derive the conditional posterior distributions of $\sigma_\star^2$, $\star\in\{R,B\}$ and σ_ϵ^2, we utilize the property that a half Cauchy random variate $U \sim \mathcal{C}^+(0,s)$ is expressed in a mixture form:

$$U^2|V \sim \mathcal{IG}\left(\frac{1}{2},\frac{1}{V}\right), \quad V \sim \mathcal{IG}\left(\frac{1}{2},\frac{1}{s^2}\right), \qquad (19)$$

where $\mathcal{IG}(a,b)$ stands for the inverse gamma distribution:

$$p(x|a,b) = \frac{b^a}{\Gamma(a)}x^{-(a+1)}e^{-\frac{b}{x}}. \qquad (20)$$

See Wand, Ormerod, Padoan, and Frühwirth (2011) and Makalic and Schmidt (2016) for more details. By introducing a latent variable $\xi_\star$, the half Cauchy distribution in (11) is rearranged as

$$\sigma_\star^2|\xi_\star \sim \mathcal{IG}\left(\frac{1}{2},\frac{1}{\xi_\star}\right), \quad \xi_\star \sim \mathcal{IG}\left(\frac{1}{2},\frac{1}{s_\star^2}\right). \qquad (21)$$

Given $\xi_\star$, we can derive the conditional posterior distribution of $\sigma_\star^2$ from (9) and (21) as

$$\begin{aligned}
p(\sigma_\star^2|\mathcal{D},\boldsymbol{\theta}_{-\sigma_\star^2},\xi_\star) &\propto p(\boldsymbol{\alpha}_\star|\mu_\star,\sigma_\star^2)p(\sigma_\star^2|\xi_\star) \\
&\propto (\sigma_\star^2)^{-\frac{N_\star}{2}}\exp\left[-\frac{\sum_{i=1}^{N_\star}(\alpha_{\star i}-\mu_\star)^2}{2\sigma_\star^2}\right]\times(\sigma_\star^2)^{-\left(\frac{1}{2}+1\right)}\exp\left(-\frac{1}{\xi_\star\sigma_\star^2}\right) \\
&\propto (\sigma_\star^2)^{-\left(\frac{N_\star+1}{2}+1\right)}\exp\left[-\frac{\frac{1}{2}\sum_{i=1}^{N_\star}(\alpha_{\star i}-\mu_\star)^2+\xi_\star^{-1}}{\sigma_\star^2}\right], \qquad (22)
\end{aligned}$$

that is,

$$\sigma_\star^2|\mathcal{D},\boldsymbol{\theta}_{-\sigma_\star},\xi_\star \sim \mathcal{IG}\left(\frac{N_\star+1}{2},\frac{\sum_{i=1}^{N_\star}(\alpha_{\star i}-\mu_\star)^2}{2}+\frac{1}{\xi_\star}\right). \tag{23}$$

Given $\sigma_\star^2$, on the other hand, the conditional posterior distribution of $\xi_\star$ is derived as

$$\begin{aligned} p(\xi_\star|\sigma_\star^2) &\propto p(\sigma_\star^2|\xi_\star)p(\xi_\star|s_\star^2) \\ &\propto \xi_\star^{-\frac{1}{2}}(\sigma_\star^2)^{-\left(\frac{1}{2}+1\right)}\exp\left(-\frac{1}{\xi_\star\sigma_\star^2}\right)\times\xi_\star^{-\left(\frac{1}{2}+1\right)}\exp\left(-\frac{1}{\xi_\star s_\star^2}\right) \\ &\propto \xi_\star^{-(1+1)}\exp\left(-\frac{\sigma_\star^{-2}+s_\star^{-2}}{\xi_\star}\right), \end{aligned} \tag{24}$$

which is the inverse gamma distribution:

$$\xi_\star|\sigma_\star^2 \sim \mathcal{IG}\left(1,\frac{1}{\sigma_\star^2}+\frac{1}{s_\star^2}\right). \tag{25}$$

Finally, we derive the conditional posterior distribution of σ_ϵ^2 and ξ_ϵ. With the mixture form of a half Cauchy distribution (19), we can rearrange (7) as

$$\sigma_\epsilon^2|\xi_\epsilon \sim \mathcal{IG}\left(\frac{1}{2},\frac{1}{\xi_\epsilon}\right),\quad \xi_\epsilon \sim \mathcal{IG}\left(\frac{1}{2},\frac{1}{s_\epsilon^2}\right), \tag{26}$$

where ξ_ϵ is a latent variable. In the same manner as (23), we can derive the conditional posterior distribution of σ_ϵ^2 from (5) and (26) as

$$\begin{aligned} p(\sigma_\epsilon^2|\mathcal{D},\boldsymbol{\theta}_{-\sigma_\epsilon^2}) &\propto p(\boldsymbol{y}|\boldsymbol{Z},\boldsymbol{\delta},\sigma_\epsilon^2)p(\sigma_\epsilon^2|\xi_\epsilon) \\ &\propto (\sigma_\epsilon^2)^{-\frac{N}{2}}\exp\left(-\frac{\sum_{i=1}^N e_i^2}{2\sigma_\epsilon^2}\right)\times(\sigma_\epsilon^2)^{-\left(\frac{1}{2}+1\right)}\exp\left(-\frac{1}{\xi_\epsilon\sigma_\epsilon^2}\right) \\ &\propto (\sigma_\epsilon^2)^{-\left(\frac{N+1}{2}+1\right)}\exp\left(-\frac{\frac{1}{2}\sum_{i=1}^N e_i^2+\xi_\epsilon^{-1}}{\sigma_\epsilon^2}\right), \end{aligned} \tag{27}$$

which is the inverse gamma distribution:

$$\sigma_\epsilon^2|\mathcal{D},\boldsymbol{\theta}_{-\sigma_\epsilon},\xi_\epsilon \sim \mathcal{IG}\left(\frac{N+1}{2},\frac{\sum_{i=1}^N e_i^2}{2}+\frac{1}{\xi_\epsilon}\right). \tag{28}$$

By replacing $\sigma_\star^2$, $\xi_\star$ and $s_\star^2$ with respectively σ_ϵ^2, ξ_ϵ and s_ϵ^2 in the derivation of (25), we obtain the conditional posterior distribution of ξ_ϵ as

$$\xi_\epsilon|\sigma_\epsilon^2 \sim \mathcal{IG}\left(1,\frac{1}{\sigma_\epsilon^2}+\frac{1}{s_\epsilon^2}\right). \tag{29}$$

Since all conditional posterior distributions (16), (18), (23), (24), (28) and (29) are standard ones, it is straightforward to set up the Gibbs sampler for generating $\boldsymbol{\theta}$ from the posterior distribution (13). However, it turns out that the plain vanilla Gibbs sampler tends to produce highly correlated sample paths which lead to inefficient estimation of parameters. In order to improve the efficiency of random number generation in the Gibbs sampler, we apply an ancillarity-sufficiency interweaving strategy (ASIS) by Yu and Meng (2011).

For this purpose, we treat $\{\alpha_{\star i}\}_{i=1}^{N_\star}$, $\star \in \{R, B\}$, as latent variables and introduce the following transformation:

$$\begin{aligned} \tilde{\alpha}_{\star i} &= \alpha_{\star i} - \mu_\star, \quad i \in \{1, \dots, N\}, \\ \tilde{y}_i &= y_i - \sum_{j=1}^{N_R} d_{Rj}^{(i)} \tilde{\alpha}_{Rj} - \sum_{j=1}^{N_B} d_{Bj}^{(i)} \tilde{\alpha}_{Bj}. \end{aligned} \tag{30}$$

Then we can rewrite the regression model (1) as

$$\tilde{y}_i = \mu_R + \mu_B \sum_{j=1}^{N_B} d_{Bj}^{(i)} + \sum_{k=1}^{K} x_{ik}\beta_k + \epsilon_{it}, \quad \epsilon_{it} \sim \mathcal{N}(0, \sigma_\epsilon^2), \tag{31}$$

because $\sum_{j=1}^{N_R} d_{Rj}^{(i)} = 1$ holds for any $i \in \{1, \dots, N\}$. Note that $\sum_{j=1}^{N_B} d_{Bj}^{(i)} = 0$ if product i is the base brand; otherwise $\sum_{j=1}^{N_B} d_{Bj}^{(i)} = 1$. The basic idea behind ASIS is that the efficiency of the Gibbs sampler depends on which specification (1) or (31) we use but it is not clear which one is better in practice. Yu and Meng (2011) proposed to combine two equivalent Gibbs samplers to improve the efficiency of the sampling algorithm.

In order to construct the ASIS algorithm, let us derive the conditional posterior distributions of the parameters in (31). By defining

$$\tilde{\boldsymbol{y}} = \begin{bmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_N \end{bmatrix}, \quad \tilde{\boldsymbol{D}} = \begin{bmatrix} 1 & \sum_{j=1}^{N_B} d_{Bj}^{(1)} \\ \vdots & \vdots \\ 1 & \sum_{j=1}^{N_B} d_{Bj}^{(N)} \end{bmatrix}, \quad \tilde{\boldsymbol{Z}} = \begin{bmatrix} \tilde{\boldsymbol{D}} & \boldsymbol{X} \end{bmatrix}, \quad \tilde{\boldsymbol{\delta}} = \begin{bmatrix} \mu_R \\ \mu_B \\ \boldsymbol{\beta} \end{bmatrix},$$

we have

$$\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{Z}}\tilde{\boldsymbol{\delta}} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}_N, \sigma_\epsilon^2 \boldsymbol{I}_N). \tag{32}$$

From (9) and (11), we obtain the prior distribution of $\tilde{\boldsymbol{\delta}}$ as

$$\tilde{\boldsymbol{\delta}} \sim \mathcal{N}\left(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}\right), \quad \tilde{\boldsymbol{\mu}} = \begin{bmatrix} \varphi_R \\ \varphi_B \\ \boldsymbol{\mu}_\beta \end{bmatrix}, \quad \tilde{\boldsymbol{\Sigma}} = \begin{bmatrix} \tau_R^2 & & \\ & \tau_B^2 & \\ & & \boldsymbol{\Sigma}_\beta \end{bmatrix}. \tag{33}$$

In the same manner as (14), we can derive the conditional posterior distribution of $\tilde{\boldsymbol{\delta}}$ from likelihood (4) and the prior distribution (33) as

$$\begin{aligned} p(\tilde{\boldsymbol{\delta}}|\mathcal{D}, \boldsymbol{\theta}_{-\tilde{\boldsymbol{\delta}}}) &\propto p(\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{Z}}, \tilde{\boldsymbol{\delta}}, \sigma_\epsilon)p(\tilde{\boldsymbol{\delta}}|\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}) \\ &\propto \exp\left[-\frac{1}{2}\left\{\sigma_\epsilon^{-2}(\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{Z}}\tilde{\boldsymbol{\delta}})^\intercal(\tilde{\boldsymbol{y}} - \tilde{\boldsymbol{Z}}\boldsymbol{\delta}) + (\tilde{\boldsymbol{\delta}} - \tilde{\boldsymbol{\mu}})^\intercal\tilde{\boldsymbol{\Sigma}}^{-1}(\tilde{\boldsymbol{\delta}} - \tilde{\boldsymbol{\mu}})\right\}\right]. \end{aligned} \tag{34}$$

By completing the square, (34) is rewritten as

$$\begin{aligned} p(\tilde{\boldsymbol{\delta}}|\mathcal{D}, \boldsymbol{\theta}_{-\tilde{\boldsymbol{\delta}}}) \propto \exp\Bigg[-\frac{1}{2}\Bigg(\tilde{\boldsymbol{\delta}} - \left(\sigma_\epsilon^{-2}\tilde{\boldsymbol{Z}}^\intercal\tilde{\boldsymbol{Z}} + \tilde{\boldsymbol{\Sigma}}^{-1}\right)^{-1}\left(\sigma_\epsilon^{-2}\tilde{\boldsymbol{Z}}^\intercal\tilde{\boldsymbol{y}} + \tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\mu}}\right)\Bigg)\left(\sigma_\epsilon^{-2}\tilde{\boldsymbol{Z}}^\intercal\tilde{\boldsymbol{Z}} + \tilde{\boldsymbol{\Sigma}}^{-1}\right) \\ \times\left(\tilde{\boldsymbol{\delta}} - \left(\sigma_\epsilon^{-2}\tilde{\boldsymbol{Z}}^\intercal\tilde{\boldsymbol{Z}} + \tilde{\boldsymbol{\Sigma}}^{-1}\right)^{-1}\left(\sigma_\epsilon^{-2}\tilde{\boldsymbol{Z}}^\intercal\tilde{\boldsymbol{y}} + \tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\mu}}\right)\right)\Bigg]. \end{aligned} \tag{35}$$

From (35), we derive the conditional posterior distribution of $\tilde{\boldsymbol{\delta}}$ as

$$\tilde{\boldsymbol{\delta}}|\mathcal{D}, \boldsymbol{\theta}_{-\boldsymbol{\delta}} \sim \mathcal{N}\left(\left(\sigma_\epsilon^{-2}\tilde{\boldsymbol{Z}}^\intercal\tilde{\boldsymbol{Z}} + \tilde{\boldsymbol{\Sigma}}^{-1}\right)^{-1}\left(\sigma_\epsilon^{-2}\tilde{\boldsymbol{Z}}^\intercal\tilde{\boldsymbol{y}} + \tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{\mu}}\right), \left(\sigma_\epsilon^{-2}\tilde{\boldsymbol{Z}}^\intercal\tilde{\boldsymbol{Z}} + \tilde{\boldsymbol{\Sigma}}^{-1}\right)^{-1}\right). \tag{36}$$

Note that, except for $\tilde{\boldsymbol{\delta}}$, the conditional posterior distributions for the rest of the parameters are the same as in (1). Thus the ASIS algorithm is given as follows.

ASIS for the hierarchical Bayes regression model

Suppose $\boldsymbol{\theta}^{(r)}$ is the r-th draw of $\boldsymbol{\theta}$ in the ASIS algorithm.

Step 1 Given $\boldsymbol{\theta}^{(r)}$, draw $(\mu_R^{(r+0.5)}, \sigma_R^{(r+0.5)}, \mu_B^{(r+0.5)}, \sigma_B^{(r+0.5)}, \boldsymbol{\beta}^{(r+0.5)}, \sigma_\epsilon^{(r+0.5)})$ via the Gibbs sampler with the conditional posterior distributions (16), (18), (23), (24), (28) and (29), and compute

$$\tilde{\alpha}_{\star i}^{(r+0.5)} = \alpha_{\star i}^{(r+0.5)} - \mu_\star^{(r+0.5)}, \quad i \in \{1, \ldots, N\}, \quad \star \in \{R, B\},$$

and obtain $\boldsymbol{\theta}^{(r+0.5)}$.

Step 2 Given $\boldsymbol{\theta}^{(r+0.5)}$, draw $(\mu_R^{(r+1)}, \sigma_R^{(r+1)}, \mu_B^{(r+1)}, \sigma_B^{(r+1)}, \boldsymbol{\beta}^{(r+1)}, \sigma_\epsilon^{(r+1)})$ via the Gibbs sampler with the conditional posterior distributions (23), (24), (28), (29) and (36), and compute

$$\alpha_{\star i}^{(r+1)} = \tilde{\alpha}_{\star i}^{(r+0.5)} + \mu_\star^{(r+1)}, \quad i \in \{1, \ldots, N\}, \quad \star \in \{R, B\},$$

and obtain $\boldsymbol{\theta}^{(r+1)}$.

Step 1 is the Gibbs sampler based on (1) while **Step 2** is the alternative sampler based on (31). The above ASIS algorithm uses two equivalent samplers in tandem so that the efficiency of random number generation will be improved.