A Bayesian analysis of vegetable production in Japan

著者	Hibiki Akira, Miyawaki Koji
journal or publication title	TUPD Discussion Papers
number	35
page range	1-32
year	2023-05
URL	http://hdl.handle.net/10097/00137190

Tohoku University Research Center for Policy Design Discussion Paper

TUPD-2023-004

A Bayesian analysis of vegetable production in Japan

Akira Hibiki
Graduate School of Economics and Management, Tohoku University

Koji Miyawaki
School of Economics, Kwansei Gakuin University
Graduate School of Economics and Management, Tohoku University

May 2023

A Bayesian analysis of vegetable production in Japan

Akira Hibiki
Research Center for Policy Design, Graduate School of Economics and Management, Tohoku University

Koji Miyawaki
School of Economics, Kwansei Gakuin University
and
Research Center for Policy Design, Graduate School of Economics and Management, Tohoku University

May 8, 2023

Abstract

Microeconomic theory often assumes that a producer maximizes its profit. As a consequence, under perfect competition, the optimal production amount is either zero or positive, where the latter satisfies the condition that the price is equal to the cost for the additional production amount (the marginal cost). This paper proposes two statistical models directly derived from this relationship and develops a Bayesian estimation method for the parameters included in this relationship. The models are applied to analyze vegetable production in Japan.

Keywords: producer theory; Bayesian approach; Jeffreys' prior.

1 Introduction

Microeconomic theory for producers starts from perfect competition, where each producer chooses its production amount given the price by maximizing its profit. When this situation as well as the single-output technology assumption holds, the optimal production amount is derived by solving the profit maximization condition and either is zero or has a positive value that satisfies the equation between the price and the production amount, where the price is equal to the change in cost from the additional production, i.e., the marginal cost. When the marginal cost is high relative to the price, a producer decides that no production amount is optimal. This paper proposes two statistical models that describe such behavior, develops an estimation methodology, and applies the models to the analysis of vegetable production in Japan.

The statistical models that are directly derived from the profit maximization condition are different from the typical regression model in two aspects. First, no production amount is explicitly included in the model. This is justified not only from a theoretical point of view, as explained above, but also from an empirical point of view. No production occurs for a substantial number of observations, as shown in the next section.

Second, the conditional distribution will be specified for the variable of the marginal cost function. We are able to transform the model into its typical form. In the econometrics literature, the former is called the structural model, while the latter is called the reduced form (see, e.g., Cameron and Trivedi (2005)). Although the parameters included in these two forms are consistent with each other, the least squares estimation does not retain this consistency. That is, the least squares estimates for the former do not yield the estimates for the latter when we use the relationship suggested by these two forms. To guarantee such a parametrization invariance as well as to address the corner solution problem, this paper takes the Bayesian approach using the Jeffreys prior.

The Jeffreys prior was proposed by Jeffreys (1946) in search for parametrization invariance and is widely known in Bayesian statistics for its desirable properties, including this invariance and the ignorance of knowledge (see Kass and Wasserman (1996)). Although this prior possesses these plausible properties, the resulting posterior distribution in our work is difficult to tract analytically, and we take the Markov chain Monte Carlo (MCMC) method to approximately infer the posterior distribution (see, e.g., Gamerman and Lopes (2006) for this method).

The proposed framework is applied to analyze vegetable production in Japan. There are 38 vegetables in our dataset, and none are produced in all prefectures of Japan. Such a production pattern is possibly due to climatic conditions (temperature and rainfall), in addition to market conditions. Then, if the marginal cost became low relative to the price due to global warming, it would be possible for producers to plant vegetables that are not produced before, as suggested by the profit maximization condition. This aspect was first examined by Mendelsohn et al. (1994) to show the impact of global warming on agriculture in the United States. It is also measured by Deschênes and Greenstone (2007).

This paper is organized as follows. After the motivating dataset is examined in Section 2, Section 3 proposes two statistical models that are directly derived from the profit maximization condition, and Section 4 describes the Bayesian estimation method. The proposed models and method are applied to analyze vegetable production in Japan in Section 5. Section 6 concludes this paper.

2 Vegetable production in Japan

The motivating data are the vegetable production data in Japan aggregated at the prefecture level in 2018, which will be analyzed in Section 5. The data are collected by the Ministry of Agriculture, Forestry and Fisheries, Japan, and consist of 47 prefectures, each of which contains about five thousand to about forty-five thousand agricultural units, about twenty thousand units on average, according to the 2020 Census of Agriculture and Forestry. Each prefecture records the production amount of 38 vegetables. This amount is different from the one sold in the market. Of $1,786$ $(= 47 \times 38)$ observations, 900 reports no production amount. Table 1 summarizes their distribution. From this table, no production is observed in about 24 prefectures on average for each vegetable.

Table 1: Distribution observations of no production amount

	Minimum	First quartile	Mean	Third quartile	Maximum
By vegetables	2	18.75	23.68	32.75	40
By prefectures	4	15	19.15	24	31

Such observations of no production occur partly because of the producer's response to climatic and market conditions. If the price were sufficiently high after taking account of these costs, all types of vegetables would be produced in positive amounts. This aspect is incorporated by resorting to the producer's profit maximization problem derived from microeconomic theory. However, its direct application causes a statistical invariance problem, as discussed in the following two sections.

3 Profit maximization condition and statistical models

Two statistical models will be described in this section. Both of them are derived from the profit maximization condition but based on different specifications of the cost function. These models are relatively simple and they ignore several issues. They will be discussed later in this section.

The first model is given by

$$\begin{cases} p_i + \lambda_i = \beta_0 + \boldsymbol{x}_i'\boldsymbol{\beta} + u_i, & \text{if } y_i = 0, \\ p_i = \alpha y_i + \beta_0 + \boldsymbol{x}_i'\boldsymbol{\beta} + u_i, & \text{if } y_i > 0, \end{cases} \tag{1}$$

where subscript i denotes the i-th observational unit ($i = 1, \ldots, n$). In the empirical analysis given in Section 5, the observational unit is the prefecture and $n = 47$ given each vegetable. In this model, y_i is the production amount, p_i is the price, and $\boldsymbol{x}_i$ is the k-dimensional covariate vector. The error term u_i is assumed to follow a normal distribution with mean 0 and variance σ^2, i.e., $N(0, \sigma^2)$.

The model parameters are $(\alpha, \beta_0, \boldsymbol{\beta}, \sigma^2, \{\lambda_i\}_{i \in \mathcal{C}_0})$, where $\mathcal{C}_0$ is a set of observational units whose y_i is zero, i.e., $\mathcal{C}_0 = \{i \mid i = 1, \ldots, n \text{ and } y_i = 0\}$. The parameter λ_i is the premium additive to the price for a producer to start production. The parameter σ^2 is the nuisance parameter, while the other parameters are the parameters of interest, i.e., the structural parameters. We assume both α and λ_i to be positive. This assumption arises naturally from the theory related to our statistical model, which is explained later, and is incorporated as prior knowledge.

This statistical model is structural because it is directly derived from microeconomic theory. In the theory, under perfect competition, given the price

p, a producer with a single-output technology chooses the production amount y that maximizes its profit subject to the following nonnegative constraint:

$$\max_y p \cdot y - C(y), \quad \text{subject to } y \geq 0,$$

where $C(y)$ is the cost function. Then, its first-order condition is given by

$$\begin{cases} p + \lambda = C', & \text{if } y = 0, \\ p = C', & \text{if } y > 0, \end{cases} \tag{2}$$

where C' is the derivative with respect to y (the so-called marginal cost in economics) and $\lambda > 0$ is the Lagrange multiplier. This is the direct result of the Karush-Kuhn-Tucker theorem. While Hall (1988) empirically suggests that this condition does not hold in the U.S. industry, no structural alternative seems to be plausible for our motivating dataset. Thus, we take it as the foundation of our analysis.

In this condition, it is the production amount that is decided. The Lagrange multiplier is sometimes called the shadow price in microeconomic theory. Under some regularity conditions, it is equal to the infinitesimal change in the nonnegative constraint. See, e.g., Mas-Colell et al. (1995) for detailed explanations of the shadow price and the producer's theory. By assuming that the marginal cost function (C') is linear in the production amount and is observed with an additive normal error, we have statistical model (1) as a natural consequence of the microeconomic theory for producers.

The relationship between the microeconomic theory provided above and model (1) implies the four following points. First, because the right-hand side of (1) is the marginal cost function, y_i and $\boldsymbol{x}_i$ are variables that determine the marginal cost. Second, the positive constraints on α and λ_i are reasonable because the marginal cost usually increases as the production amount increases and because no production amount occurs when the marginal cost is more than the price, respectively.

Third, the error term represents variations in the omitted variables and the deviation from the first-order approximation of the marginal cost function (i.e., the second- and higher-order terms). It would be more rigorous to truncate the distribution because it is the difference between the price and the specified marginal cost function. Regarding this point, we note that the next model (4) addresses this issue, and we provide a more detailed discussion below.

Fourth, the dependent variable is not p_i but y_i. This aspect requires special care when we estimate the model parameters. A similar and familiar form is the Tobit regression model, where y_i is on its left hand, which is given by

$$\begin{aligned} y_i &= \max\{y_i^*, 0\}, \\ y_i^* &= \delta_0 + \gamma p_i + \boldsymbol{x}_i' \boldsymbol{\delta} + v_i, \end{aligned} \tag{3}$$

where y_i^* is the latent production amount and the error term v_i is assumed to follow a normal distribution with mean 0 and variance τ^2 (see Chib (1992) for its Bayesian estimation method). Both model (1) and model (3) should be consistent with their theoretical origin, model (2). That is, the model parameters $(\alpha, \beta_0, \boldsymbol{\beta}, \sigma^2, \{\lambda_i\})$ and $(\gamma, \delta_0, \boldsymbol{\delta}, \tau^2, \{y_i^*\})$ are consistent with each other. Such consistency is easier to implement in the Bayesian estimation. Section 4 discusses this aspect further.

The second model is given by

$$\begin{cases} \log p_i + \lambda_i = \beta_0 + \boldsymbol{x}_i' \boldsymbol{\beta} + u_i, & \text{if } y_i = 0, \\ \log p_i = \alpha \log y_i + \beta_0 + \boldsymbol{x}_i' \boldsymbol{\beta} + u_i, & \text{if } y_i > 1. \end{cases} \tag{4}$$

Because this model is essentially the same from a statistical point of view, we reuse the parameters. However, their interpretation is different, e.g., λ_i is not the additive premium but the logarithm of the multiplicative premium to start production. In addition, the production amount jumps from $y_i = 0$ to $y_i = 1$ as the price rises. Then, the term on the left-hand side, $\beta_0 + \boldsymbol{x}_i'\boldsymbol{\beta} + u_i$ if $y_i = 0$, becomes the logarithm of the so-called shutdown price in economics. A possible source of this jump is the nonsunk fixed cost to produce the amount of the first unit (see Mas-Colell et al. (1995)). From the perspective of microeconomic theory, the log model (4) is more reasonable in agriculture than the linear model (1) because agriculture usually requires the rental and depreciation costs regardless of the production amount.

These two models have advantages and weak points. The linear model is a direct derivation from the profit maximization condition and puts no constraints on the measurement unit. However, it requires a truncation, such as $I(p_i - \beta_0 - \boldsymbol{\beta}'\boldsymbol{x}_i > u_i)$ for $y_i > 0$, as discussed above. A similar truncation occurs under the reduced form, which leads to the complicated Jeffreys prior, as noted in Chib (1992) and Amemiya (1973). Further, Angrist (2001) discusses that the marginal effect without truncation on the dependent variable is estimated to be similar to the marginal effect with truncation. Therefore, we do not incorporate this truncation in the linear model to avoid such a complication.

The log model has a jump, as described above, and requires no truncation of the error term. However, it requires a constraint on the measurement unit. The unit should be selected as long as it satisfies $y_i > 1$ for all observations with a positive production amount. This holds in our empirical dataset, the minimum of which is 47.

The linear and log models are different with respect to the cost function that they are based on. When the cost function is specified as a quadratic function of the production amount, we obtain the linear model.

When $\log C(y) = a + b \log(y)$, the marginal cost function is $C' = bC(y)/y$, leading to

$$\log C' = a + \log(b) + (b - 1) \log(y).$$

Thus, it is straightforward to obtain the log model (4) with this specific cost function. Model selection between these two models by the deviance information criterion (DIC) proposed by Spiegelhalter et al. (2002) will be conducted in Subsection 5.2 with the empirical dataset.

Four issues related to the proposed statistical models are discussed. First, the decision on the production amount is usually made before the actual equilibrium is observed, and it might be more reasonable to use the expectation of the price instead of its current values (the price contemporaneous with the production amount), provided there is no uncertainty about the production amount. A possible explanation for this situation is the rational expectation assumption where expected values are observed.

This paper assumes the rational expectation, mainly because Cooley and DeCanio (1977) and Goodwin and Sheffrin (1982) empirically showed that this assumption is consistent with agricultural data. In addition, Subsection 5.4 uses the previous price instead of the current one and shows that this change has little influence on the inferences about the parameters of interest.

The decision under uncertainty would be another aspect to consider. Sandmo (1971) proposes its analytical framework, while Pope and Chavas (1994), Pope and Just (1996), and Moschini (2001) provide its empirical specifications. However, to keep the models as simple as possible and to avoid misspecifying the structure based on expectation, we do not pursue this aspect in this paper.

Second, model parameter estimates can suffer from the endogeneity bias. As an example, consider a factor that relates to the economic environment. It may also relate to the producer's technology, so that the error term correlates with the factor that relates to the economic environment if such a factor is excluded from the model. On the other hand, this factor also affects the demand, so that the price is likely to correlate with this factor. Then, this factor omitted from the model causes the endogeneity bias on model parameter estimates. When we assume that the technology is independent of such a factor, it is not a problem. The short-run analysis is a case where this assumption holds because the technology does not change so quickly.

Third, the marginal cost function is specified as linear in the production amount, even after taking logarithms. An alternative is its nonparametric estimation proposed by Hall and Yatchew (2007). Our specification can be considered to be its first-order approximation. Estimation with higher-order terms complicates the estimation methodology, and we leave it for a future work.

Finally, the substitution between vegetables is ignored in this model. Thus, as the theoretical foundation, we assume a producer with a single-output technology. Depending on the application context, the substitution would be of interest. A theoretical framework for this case is the one proposed by Pfouts (1961), where a firm produces many products. Such a situation is observed in the manufacturing industry, and is examined by Bernard et al. (2010) with a continuum of products. A possible empirical specification is proposed by Just et al. (1983), useful for the inner solution.

In our model, such an aspect can be incorporated when the land is included in the profit maximization condition. We do not include it by assuming that the producer has the unlimited access to the land. As in our empirical analysis, this assumption is reasonable when the production unit is the prefecture, which can be considered to have enough land.

4 Bayesian estimation

This section describes the Bayesian estimation of the linear model (1). The estimation of the log model (4) can be implemented in a similar way, which will not be described in this paper (see Remark 1 below).

The likelihood function under the linear model (1) is proportional to

$$f\left(\{y_i\}_{i=1}^n \mid \alpha, \beta_0, \boldsymbol{\beta}, \sigma^2, \{\lambda_i\}_{i\in\mathcal{C}_0}, \{p_i, \boldsymbol{x}_i\}_{i=1}^n\right) \propto |\alpha|^{n_1}\left(\sigma^2\right)^{-n/2}$$
$$\times \exp\left[-\frac{1}{2\sigma^2}\left\{\sum_{i\in\mathcal{C}_0}\left(p_i+\lambda_i-\beta_0-\boldsymbol{x}_i'\boldsymbol{\beta}\right)^2+\sum_{i\in\mathcal{C}_1}\left(p_i-\alpha y_i-\beta_0-\boldsymbol{x}_i'\boldsymbol{\beta}\right)^2\right\}\right],$$

where $\mathcal{C}_1$ is the complement of $\mathcal{C}_0$, i.e., $\mathcal{C}_1 = \{i \mid i = 1, \ldots, n \text{ and } y_i > 0\}$ and $n_1 = |\mathcal{C}_1|$. The design matrices $\boldsymbol{Z}$, $\boldsymbol{X}_1$, and $\boldsymbol{X}_2$ are assumed to be of full column rank, where $\boldsymbol{Z}$ has $(1, \boldsymbol{x}_i')$ for its i-th row $(i = 1, \ldots, n)$, $\boldsymbol{X}_1$ has $(p_i, y_i, 1, \boldsymbol{x}_i')$ for its i-th row $(i \in \mathcal{C}_1)$, and $\boldsymbol{X}_2$ has $(p_i, 1, \boldsymbol{x}_i')$ for its i-th row $(i \in \mathcal{C}_0)$.

Remark 1. For the log model, change (p_i, y_i) to $(\log p_i, \log y_i)$ and modify $\mathcal{C}_1 = \{i \mid i = 1, \ldots, n \text{ and } y_i > 1\}$ to have the likelihood with respect to $\{\log y_i\}_{i=1}^n$. Then, the following discussion holds similarly.

The prior density function is assumed to be

$$\pi\left(\alpha, \beta_0, \boldsymbol{\beta}, \sigma^2, \{\lambda_i\}_{i\in\mathcal{C}_0}\right) \propto \frac{1}{\alpha} I(\alpha > 0) \frac{1}{\left(\sigma^2\right)^{(k+4+n_0)/2}} I\left(\sigma^2 > 0\right)$$
$$\times \left\{\frac{1}{\sigma^2} + \frac{q_1^2}{2n_1\left(1 - \frac{n_1}{n}\right)}\right\}^{1/2} \prod_{i\in\mathcal{C}_0} I(\lambda_i > 0), \quad (5)$$

where $n_0 = |\mathcal{C}_0|$ and

$$q_1^2 = \sum_{i\in\mathcal{C}_1} p_i^2 - \sum_{i\in\mathcal{C}_1} \boldsymbol{z}_i' p_i \left(\sum_{i\in\mathcal{C}_1} \boldsymbol{z}_i \boldsymbol{z}_i'\right)^{-1} \sum_{i\in\mathcal{C}_1} \boldsymbol{z}_i p_i.$$

We use $I(A)$ as the indicator function, where $I(A) = 1$ if A is true and $I(A) = 0$ otherwise. This prior is a product of the prior knowledge of the sign constraint and the so-called Jeffreys prior derived from the structural model (1). Appendix A provides its derivation. The Jeffreys prior is often used in the objective Bayes literature and considered to be one of the noninformative priors (see, e.g., Kass and Wasserman (1996)). Because a noninformative prior does not necessarily lead to a proper posterior distribution, the proof of its properness is given in Appendix B.

One main advantage of the Jeffreys prior is its invariance to parameter transformation. The proposed statistical model has another parametrization, where y_i is on the left-hand side (3). Both parametrizations should lead to posterior distributions that are consistent with each other after the appropriate parameter transformation. Otherwise, the prior distribution under one model introduces prior knowledge that is different from that under the other model. Such invariance is a desirable property for the statistical inference of our models. This paper focuses on the parametrization of model (1) and its Jeffreys prior because its statistical inferences are easier to conduct.

Our prior specification has a side effect. Eaton and Sudderth (2010) show that the Jeffreys' prior corresponds to the left-invariant Haar measure in the invariance situation (see their Theorem 3.7). This implies that the resulting estimator is invariant according to the linear unit transformation (from ton to kilogram, for example). Because the choice of unit is arbitrary in our empirical application, this property is plausible.

To conduct inferences, we apply the Gibbs sampler (one of the Markov chain Monte Carlo methods) to approximately infer the posterior distribution. It is implemented in the following six-step algorithm.

Step 1. Initialize the model parameters $(\alpha, \beta_0, \boldsymbol{\beta}, \sigma^2, \{\lambda_i\}_{i \in \mathcal{C}_0})$.

Step 2. Generate α conditional on $\beta_0, \boldsymbol{\beta}, \sigma^2, \{\lambda_i\}_{i \in \mathcal{C}_0}$.

Step 3. Generate $(\beta_0, \boldsymbol{\beta})$ conditional on $\alpha, \sigma^2, \{\lambda_i\}_{i \in \mathcal{C}_0}$

Step 4. Generate σ^2 conditional on $\alpha, \beta_0, \boldsymbol{\beta}, \{\lambda_i\}_{i \in \mathcal{C}_0}$.

Step 5. Generate λ_i conditional on $\alpha, \beta_0, \boldsymbol{\beta}, \sigma^2$ for $i \in \mathcal{C}_0$.

Step 6. Repeat Step 2 through Step 5.

Because the conditional densities used in Steps 2 and 4 are nonstandard, we use the Metropolis-Hastings algorithm to generate a random sample from them. For Step 2, the Laplace approximation based on the Taylor series expansion of its logarithm around the mode is applied (see Tierney and Kadane (1986)). For Step 4, a mixture of inverse gamma distributions is used as the proposal distribution. For details of the full conditional posterior density functions used in this Gibbs sampler, see Appendix C.

5 Empirical analysis of Japanese vegetable production

This section analyzes the prefecture-level vegetable production data in Japan in 2018. As stated in Section 2, 38 vegetables are included in the analysis, and they are separately analyzed by applying the models described in Section 3.

5.1 Data description

The variables to be used for the analysis are the production amount, the price, and climate-related variables. See also Section 2 for the source of the dataset. The measurement unit for the production amount and the price are ton and Japanese yen per kilogram, respectively. The price variable is the one reported at the wholesale market. When a prefecture has more than one market, we average the prices by using their sales amount at markets as weight. For the climate-related variables, three variables are selected, i.e, the average temperature x_1, the average rainfall x_2, and the average number of hours of sunlight x_3, collected by the Japan Meteorological Agency. Although these variables vary even within a prefecture, we use the ones reported at the meteorological observatory closest to the prefectural capital.

Wage and the rental cost of capital would be reasonable candidate variables for the marginal cost function. However, they are not available in Japan partly because most agricultural units are family-based (about 96.4% of them are owned by an individual according to the 2020 Census of Agriculture and Forestry), so that wage and the rental cost of capital are difficult to measure for such a unit.

Because the log model (4) is preferred to the linear model (1) in terms of DIC, as shown later, the logarithm of the production amount and price are summarized in the form of a boxplot by each vegetable in Figures 1 and 2. The box gives the interquartile range (IQR) between the first and third quartiles with a line at the median. The two whiskers are the 1.5-IQR length, and observations outside the whiskers are given by dots. When observations are within the 1.5-IQR length, the whiskers are stretched at the maximum/minimum.

Summary statistics of the climate-related variables are given in Table 2.

5.2 MCMC settings and model selection

For all vegetable data, we estimate the model parameters by the MCMC method described in the previous section. Discarding the first 10,000 samples, the subsequent 100,000 samples are generated, and every 5-th sample

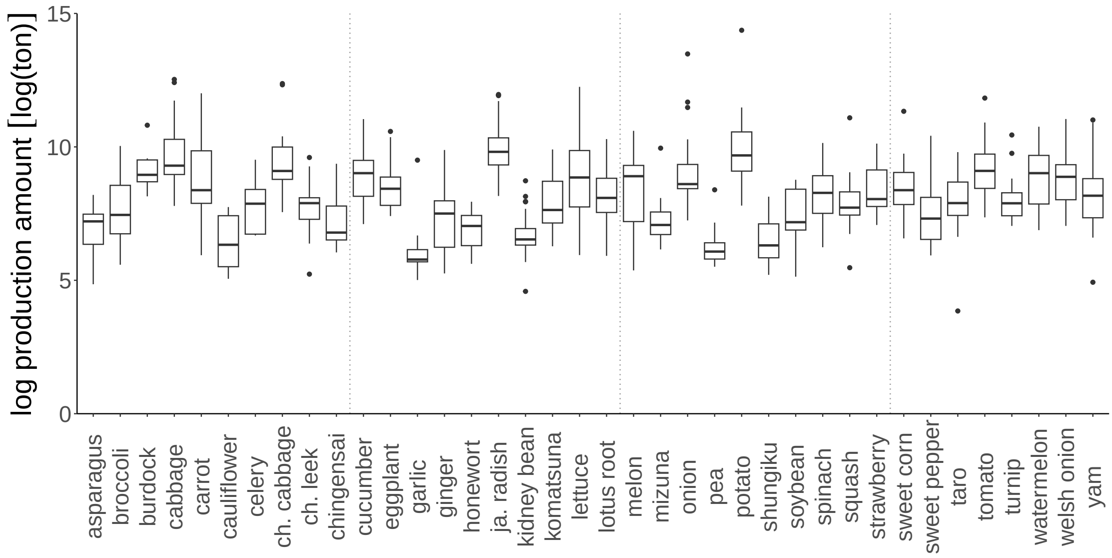

Figure 1: Log production amount by vegetables.

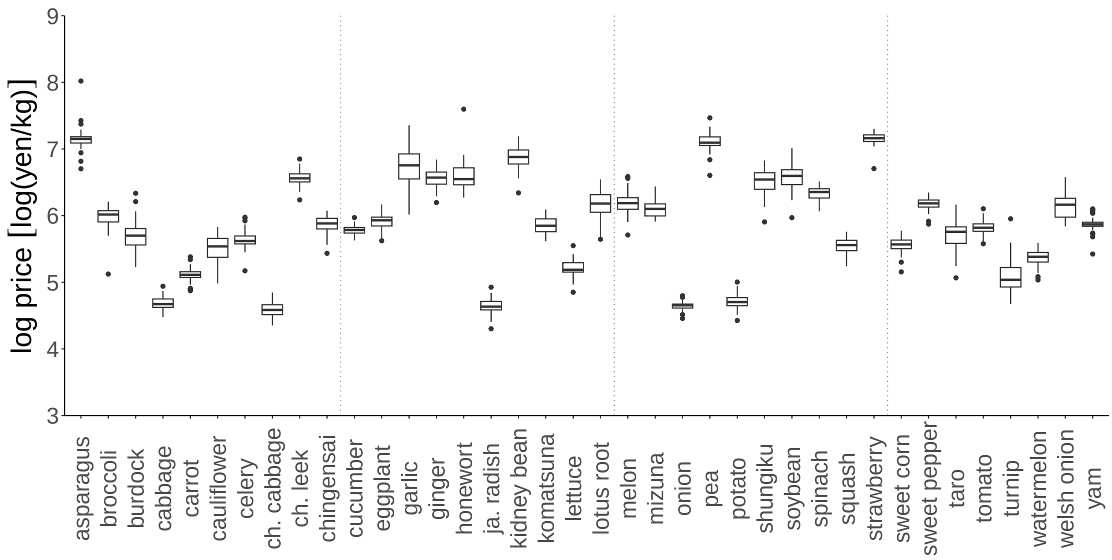

Figure 2: Log price by vegetables.

is selected to have 20,000 samples. They are used to conduct the following statistical inferences. The inefficiency factor is used to check whether the number of MCMC samples is high enough (see Chib (2004) for this measure).

Table 2: Summary statistics of the climate-related variables

	Unit	Minimum	First quartile	Mean	Third quartile	Maximum
average temperature (x_1)	°C	9.55	15.38	16.10	17.31	23.51
average rainfall (x_2)	mm	2.27	3.74	4.87	5.47	8.68
average hours of sunlight (x_3)	hour	4.18	5.20	5.67	6.13	6.55

All inefficiency factors are less than 50.

We compare the model fitting between the linear and log models by DIC. We note that the likelihood with respect to y_i is used for DIC, so that it is comparable between both models. After the MCMC samples are divided into ten equally sized groups, DIC is estimated for each group. Their average and standard deviation are an estimate of DIC and its standard error, respectively. Judging from the estimated DIC, the log model is preferred, because the minimum difference is about 18, with a standard error of about 0.43, and the other standard errors are less than 1.

5.3 Analysis of the model parameters

Figure 3 shows boxplots of the (marginal) posterior distributions of the average of $\exp(\lambda_i)$, that is, $\bar{\lambda} = \frac{1}{n_0}\sum_{i\in\mathcal{C}_0}\exp(\lambda_i)$, for all vegetables. Each box represents the first and third quartiles with a line at the mean, and both ends of the whiskers show the 95% credible interval.

As the log model is used, this average can be interpreted as the average (multiplicative) premium price to start production scaled by the price. The premium differs across vegetables, reflecting their economic and agricultural environment. Most premium prices are around 1.2, and they are less than about 1.8.

Figure 4 shows the marginal effect of the average temperature on the

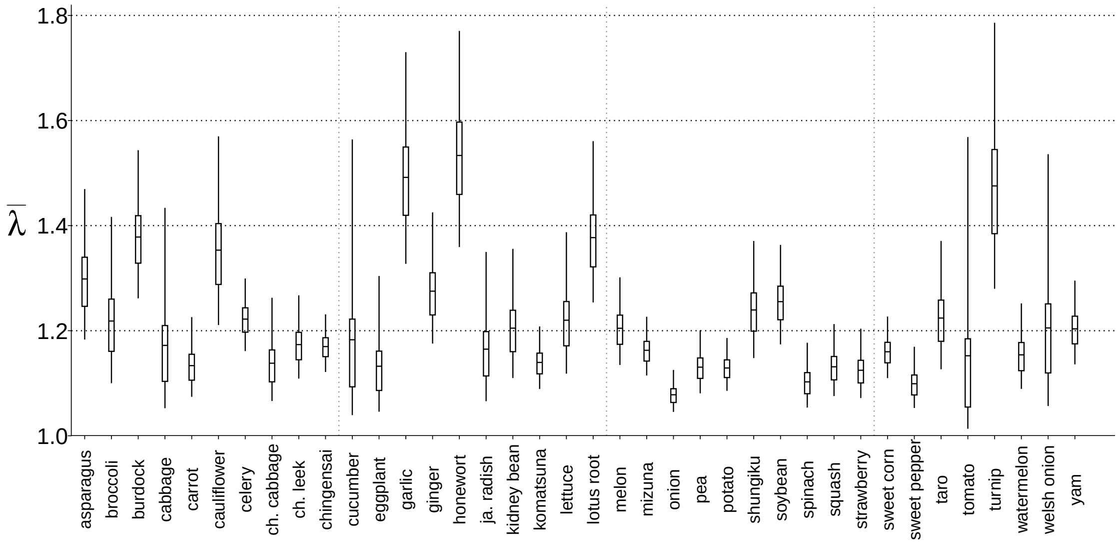

Figure 3: Premium.

marginal cost. Boxplots are drawn in the same manner as the ones in Fig-

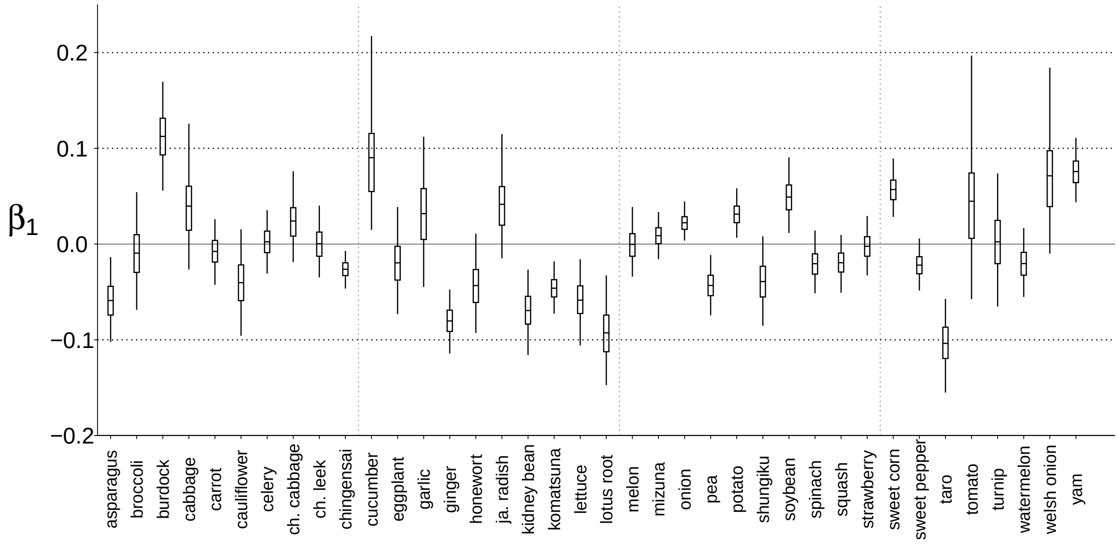

Figure 4: Marginal effect of the average temperature.

ure 3. This figure suggests that the association between the marginal cost and the average temperature is nonuniform among vegetables. In its Sixth

Assessment Report by Working Group II, the Intergovernmental Panel on Climate Change reports that the average temperature will rise by about 1.5 to 5 °C in 2100 compared with the one in 1850–1900 (see Pörtner et al. (2022)). Our result suggests that its influence on vegetable production varies according to the type of vegetable.

To see this more clearly, we divide the vegetables into two groups according to their marginal posterior probabilities of β_1. For each vegetable, we estimate its posterior probability that its β_1 is positive/negative, i.e., $\Pr(\beta_1 > 0 \mid \text{data})$ or $\Pr(\beta_1 < 0 \mid \text{data})$. When the probability is more than 0.95, it is credible that the average temperature has a positive/negative association with the marginal cost. The results are given in Table 3.

Table 3: Marginal posterior probability of β_2

Positive	Negative
burdock, cucumber, potato, welsh onion, onion, sweet corn, soybean, yam	asparagus, chingensai, ginger, kidney bean, komatsuna, lettuce, lotus root, pea, shungiku, taro

The posterior summary for the remaining model parameters is provided in Appendix D.

5.4 Examination of the rational expectation assumption

The result above relies on the rational expectation assumption, as discussed in Section 3. To examine whether this assumption is reasonable, we use the price in 2017 instead of 2018 and see how this assumption affects the result. The same prior and MCMC settings as for the log model are used to generate 20,000 samples. Figure 5 shows the comparison of posterior means with the datasets in 2017 and 2018, overlaying the 45-degree line, for all model parameters. Judging from the posterior mean, the two results are mostly similar because most points gather around the 45-degree line and no systematic deviation from the line exists. Therefore, we conclude that relying on the rational expectation assumption would be reasonable, at least for this dataset.

6 Discussion

This paper proposed two statistical models that are derived from the producer's profit maximization condition. To estimate the model parameters, we take the Bayesian approach using the Jeffreys prior to guarantee parametrization invariance between the structural and reduced forms. The proposed models and estimation method are applied to analyze vegetable production in Japan. The analysis estimates the premium price to start production and the impact of the average temperature rise on the production amount. The latter shows a nonuniform influence of global warming on vegetable production.

To analyze the impact of global warming more accurately, we need other information, such as the demand structure. In addition, as discussed in the previous section, the substitution between agricultural products is another factor to incorporate when we analyze the impact, which can be addressed by the panel dataset. Finally, relaxing the assumption of perfect competition would be useful when our framework is applied to another industry. Implementing the above points requires a complication of our methodology and

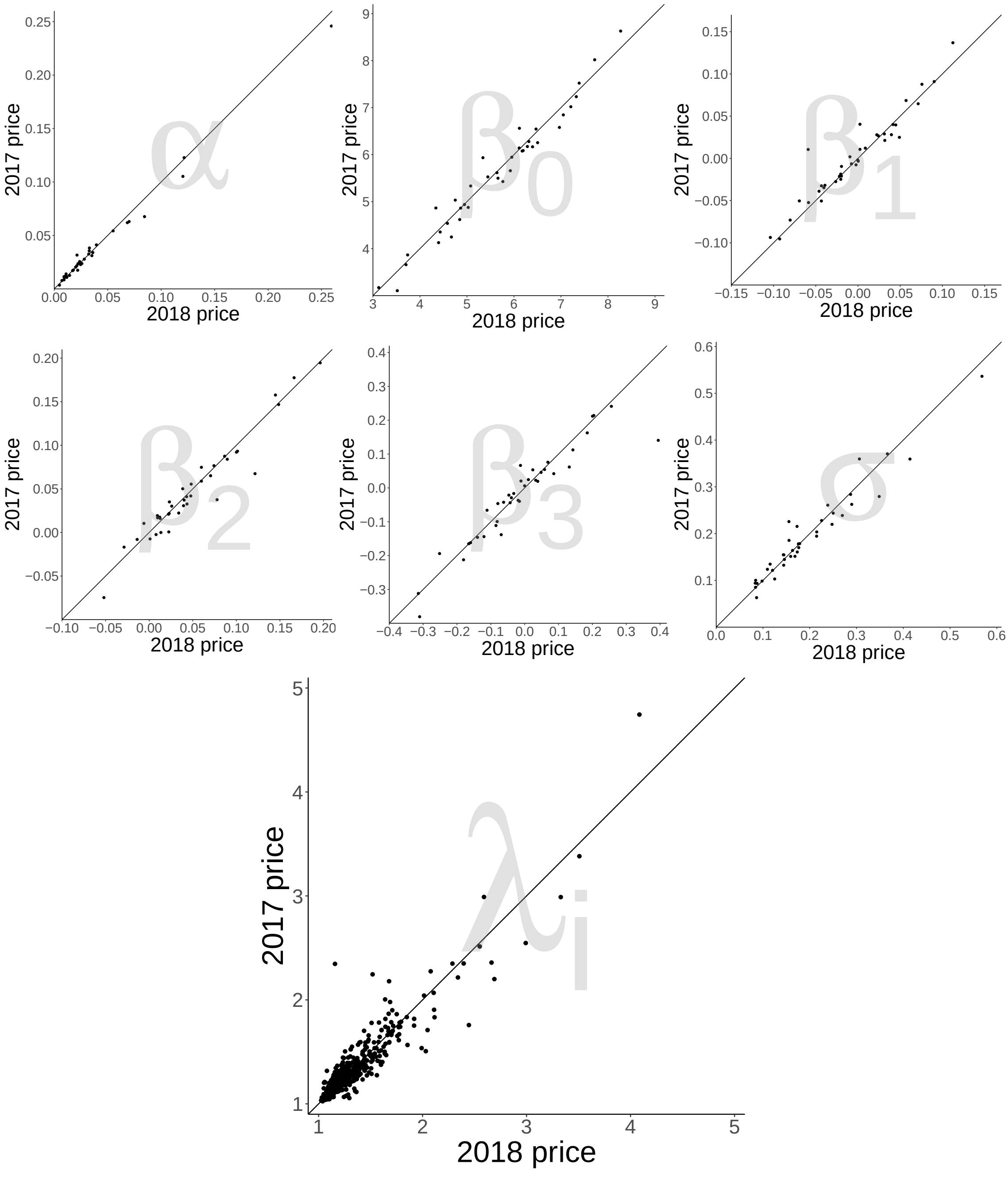

Figure 5: Comparison.

will be addressed in future works.

A Jeffreys' prior

The Jeffreys prior is proportional to the square root of the determinant of the Fisher information matrix. To this end, the following formula for the determinant of a partitioned matrix is applied: $|\boldsymbol{A}| = |\boldsymbol{A}_{11}||\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}| = |\boldsymbol{A}_{22}||\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}|$, where

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}.$$

See, e.g., Abadir and Magnus (2005) for a proof of this formula.

The Fisher information matrix for the linear model is given by

$$\boldsymbol{F} = -E \begin{pmatrix} \frac{\partial^2 \log f}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta}'} & \frac{\partial^2 \log f}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\lambda}'} \\ \frac{\partial^2 \log f}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\zeta}'} & \frac{\partial^2 \log f}{\partial \boldsymbol{\lambda} \partial \boldsymbol{\lambda}'} \end{pmatrix},$$

where $\boldsymbol{\zeta} = (\alpha, \sigma^2, \beta_0, \boldsymbol{\beta}')'$ and $\boldsymbol{\lambda} = (\lambda_i)_{i \in \mathcal{C}_0}$. The expectation is over the conditional distribution of y_i $(i = 1, \ldots, n)$. This is a partitioned matrix, where $\boldsymbol{F}_{ij}$ is its (i, j) block element for $i, j = 1, 2$. After some calculations, we have

$$\begin{aligned}
\boldsymbol{F}_{11} &= \begin{pmatrix} \frac{2n_1}{\alpha^2} + \frac{1}{\alpha^2\sigma^2}\sum_{i\in\mathcal{C}_1} d_i^2 & -\frac{n_1}{\alpha\sigma^2} & \frac{1}{\alpha\sigma^2}\sum_{i\in\mathcal{C}_1} \boldsymbol{z}_i' d_i \\ -\frac{n_1}{\alpha\sigma^2} & \frac{n}{2(\sigma^2)^2} & 0 \\ \frac{1}{\alpha\sigma^2}\sum_{i\in\mathcal{C}_1} \boldsymbol{z}_i d_i & 0 & \frac{1}{\sigma^2}\sum_i \boldsymbol{z}_i \boldsymbol{z}_i' \end{pmatrix}, \\
\boldsymbol{F}_{22} &= \frac{1}{\sigma^2}\boldsymbol{I}_{n_0}, \\
\boldsymbol{F}_{12} &= \boldsymbol{F}_{21}' = \begin{pmatrix} \boldsymbol{0}' \\ \boldsymbol{0}' \\ -\frac{1}{\sigma^2}\boldsymbol{z}_i \end{pmatrix},
\end{aligned}$$

where $\boldsymbol{z}_i' = (1, \boldsymbol{x}_i')$, $\boldsymbol{I}_{n_0}$ is the n_0-dimensional unit matrix, the index i in $\boldsymbol{F}_{12}$ is over $i \in \mathcal{C}_0$, and

$$d_i = \begin{cases} p_i + \lambda_i - \beta_0 - \boldsymbol{x}_i'\boldsymbol{\beta}, & \text{if } i \in \mathcal{C}_0, \\ p_i - \beta_0 - \boldsymbol{x}_i'\boldsymbol{\beta}, & \text{if } i \in \mathcal{C}_1. \end{cases} \tag{6}$$

Then,

$$\boldsymbol{F}_{11} - \boldsymbol{F}_{12}\boldsymbol{F}_{22}^{-1}\boldsymbol{F}_{21} = \begin{pmatrix} \frac{2n_1}{\alpha^2} + \frac{1}{\alpha^2\sigma^2}\sum_{i\in\mathcal{C}_1} d_i^2 & -\frac{n_1}{\alpha\sigma^2} & \frac{1}{\alpha\sigma^2}\sum_{i\in\mathcal{C}_1} \boldsymbol{z}_i' d_i \\ -\frac{n_1}{\alpha\sigma^2} & \frac{n}{2(\sigma^2)^2} & 0 \\ \frac{1}{\alpha\sigma^2}\sum_{i\in\mathcal{C}_1} \boldsymbol{z}_i d_i & 0 & \frac{1}{\sigma^2}\sum_{i\in\mathcal{C}_1} \boldsymbol{z}_i\boldsymbol{z}_i' \end{pmatrix}.$$

After combining all the expressions above and picking up the terms that include the model parameters, we have

$$\begin{aligned} |\boldsymbol{F}| &= \begin{vmatrix} \frac{n}{2(\sigma^2)^2} & 0 \\ 0 & \frac{1}{\sigma^2}\sum_{i\in\mathcal{C}_1} \boldsymbol{z}_i\boldsymbol{z}_i' \end{vmatrix} \times \left\{ \frac{2n_1}{\alpha^2}\left(1 - \frac{n_1}{n}\right) + \frac{1}{\alpha^2\sigma^2} q_1^2 \right\} \\ &\propto \frac{1}{\alpha^2 \left(\sigma^2\right)^{k+4+n_0}} \left\{ \sigma^2 + \frac{q_1^2}{2n_1\left(1 - \frac{n_1}{n}\right)} \right\}. \end{aligned}$$

Therefore, we have the Jeffreys prior found in Equation (5).

B Properness of the posterior distribution

This section will give a proof that the posterior distribution that uses Prior (5) is proper. The posterior density is proportional to

$$\begin{aligned}
\pi\left(\alpha, \beta_0, \boldsymbol{\beta}, \sigma^2, \{\lambda_i\}_{i\in\mathcal{C}_0}\right) \propto & |\alpha|^{n_1-1}\left(\sigma^2\right)^{-(k+n_0+n+4)/2}\exp\left\{-\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(d_i-\alpha y_i\right)^2\right\} \\
& \times I\left(\alpha>0\right) I\left(\sigma^2>0\right)\left\{\frac{1}{\sigma^2}+\frac{q_1^2}{2n_1\left(1-\frac{n_1}{n}\right)}\right\}^{1/2} \\
& \times \prod_{i\in\mathcal{C}_0} I\left(\lambda_i>0\right) \\
< & |\alpha|^{n_1-1}\left(\sigma^2\right)^{-(k+n_0+n+4)/2}\exp\left\{-\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(d_i-\alpha y_i\right)^2\right\} \\
& \times I\left(\alpha>0\right) I\left(\sigma^2>0\right)\left(\frac{1}{\sigma}+\frac{q_1}{\sqrt{2n_1\left(1-\frac{n_1}{n}\right)}}\right) \\
& \times \prod_{i\in\mathcal{C}_0} I\left(\lambda_i>0\right).
\end{aligned}$$

Our proof shows that the integral of the function on the most right-hand side over its parameter space is finite.

With the rank condition, two normalizing constants from the inverse gamma distribution and the Arellano-Valle and Bolfarine generalized t distribution are finite. Then, we have

$$\begin{aligned}
\int \pi\left(\alpha, \beta_0, \boldsymbol{\beta}, \sigma^2, \{\lambda_i\}_{i\in\mathcal{C}_0}\right) d\sigma^2 d\beta_0 d\boldsymbol{\beta} < & (\text{some finite constant}) \\
& \times |\alpha|^{n-1}\left(\boldsymbol{e}'\boldsymbol{M}\boldsymbol{e}\right)^{-(n_0+n+1)/2},
\end{aligned}$$

where $\boldsymbol{M} = \boldsymbol{I}_n - \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'$, $\boldsymbol{e}$ is the n-dimensional vector whose i-th row is defined as

$$e_i = \begin{cases} p_i + \lambda_i - \alpha y_i, & \text{if } i \in \mathcal{C}_0, \\ p_i - \alpha y_i, & \text{if } i \in \mathcal{C}_1. \end{cases} \tag{7}$$

Let

$$f\left(\alpha, \{\lambda_i\}_{i \in \mathcal{C}_0}\right) = \left(\alpha^2\right)^{(n-1)/2} (\boldsymbol{e}'\boldsymbol{M}\boldsymbol{e})^{-(n_0+n+1)/2}.$$

This function is integrable on any compact subset A of the nonnegative orthant because of the following three reasons: (i) $|\alpha|^{(n-1)/2}$ is continuous and closed on A, (ii) $\boldsymbol{e}'\boldsymbol{M}\boldsymbol{e}$ is so as well, and (iii) $\boldsymbol{e}'\boldsymbol{M}\boldsymbol{e} > 0$ because of the rank condition.

Further, let

$$g\left(\alpha, \{\lambda_i\}_{i \in \mathcal{C}_0}\right) = \left(\|\boldsymbol{\xi}\|^2\right)^{(n-1)/2} (\boldsymbol{e}'\boldsymbol{M}\boldsymbol{e})^{-(n_0+n+1)/2},$$

where $\boldsymbol{\xi} = (\alpha, \{\lambda_i\}_{i \in \mathcal{C}_0})$ and $\|\boldsymbol{\xi}\|$ is the Euclidean norm of $\boldsymbol{\xi}$. It is clear that $f \leq g$ for any possible value of $\boldsymbol{\xi}$. This function is integrable on A.

We will show that $g = O(\|\boldsymbol{\xi}\|^{-(n_0+2)})$, where the order notation denotes $|g| \leq M\|\boldsymbol{\xi}\|^{-(n_0+2)}$ for some constant M. The polar coordinate representation of $\boldsymbol{\xi}$ leads to

$$\frac{\|\boldsymbol{\xi}\|^2}{\boldsymbol{e}'\boldsymbol{M}\boldsymbol{e}} = \frac{dr^2}{a + br + cr^2},$$

where a, b, c, d are some functions of triangular functions and r is the radial distance from the origin. We note that $cr^2 = (\boldsymbol{e} - \boldsymbol{p})'\boldsymbol{M}(\boldsymbol{e} - \boldsymbol{p})$, where $\boldsymbol{p} = (p_1, \ldots, p_n)'$. By the rank condition, $cr^2 > 0$. So $c \neq 0$ as long as $r > 0$. Then, the left-hand side converges to $1/c$ as $\|\boldsymbol{\xi}\|$ goes to infinity. Because

$$g \times \left(\|\boldsymbol{\xi}\|^{(n_0+2)}\right) = \left(\frac{\|\boldsymbol{\xi}\|^2}{\boldsymbol{e}'\boldsymbol{M}\boldsymbol{e}}\right)^{(n_0+n+1)/2},$$

we have the result.

Because the order result shows that g is integrable over the nonnegative orthant, f is so as well, which completes the proof.

C Full conditional posterior density functions

First, the full conditional posterior density for α is proportional to

$$\begin{aligned}\pi\left(\alpha \mid \beta_0, \boldsymbol{\beta}, \sigma^2, \{\lambda_i\}_{i\in\mathcal{C}_0}\right) \propto & |\alpha|^{n_1-1} \\ & \times \exp\left[-\frac{1}{2\sigma^2}\left\{\alpha^2\left(\sum_{i=1}^{n} y_i^2\right) - 2\alpha\left(\sum_{i=1}^{n} y_i d_i\right)\right\}\right] \\ = & |\alpha|^{n_1-1}\exp\left(-\frac{R}{2}\alpha^2 + Q\alpha\right),\end{aligned}$$

where $Q = \sigma^{-2}\sum_{i=1}^{n} y_i d_i$ (see Equation (6) for d_i) and $R = \sigma^{-2}\sum_{i=1}^{n} y_i^2$. This conditional density is nonstandard, and we apply the Metropolis-Hastings (MH) algorithm to draw a sample from it.

The proposal used in this step is derived as follows. The mode of this conditional density is

$$m = \frac{Q + \sqrt{Q^2 + 4R\,(n_1 - 1)}}{2R},$$

and the second derivative of the log conditional density is given by

$$H = -\frac{n_1 - 1}{m^2} - R.$$

Then, the Taylor series expansion of the log conditional density around the mode gives the proposal $TN_{(0,\infty)}(m, -H^{-1})$, where $TN_S(\mu, \tau^2)$ denotes the truncated normal distribution with mean μ and variance τ^2 on support S. Let $\alpha^{(-1)}$ and $\tilde{\alpha}$ be the sample recorded in the previous Markov chain and the candidate drawn from the proposal density, respectively. The candidate is accepted with probability

$$\min\left[1, \frac{g\left(\tilde{\alpha}\right)}{g\left(\alpha^{(-1)}\right)}\right],$$

where

$$g\left(\alpha\right) = |\alpha|^{n_1-1}\exp\left\{-\frac{1}{2}\left(R + H\right)\alpha^2 + \left(Q + mH\right)\alpha\right\}.$$

Next, because the full conditional density for σ^2 is also nonstandard, the MH algorithm is applied as well. The proposal for this generation is a mixture of the inverse gamma distributions $IG(r_1/2, S_1/2)$ and $IG((r_1+1)/2, S_1/2)$ with the respective weights w and $1-w$, where $r_1 = k+2+n+n_0$, $S_1 = \sum_{i=1}^n (d_i - \alpha y_i)^2$, and

$$w = \frac{q_1}{q_1 + \sqrt{2n_1\left(1-\frac{n_1}{n}\right)}}.$$

The acceptance probability for a candidate $\tilde{\sigma}^2$ in terms of the previous sample $\sigma^{2,(-1)}$ is given by

$$\min\left[1, \frac{g\left(\tilde{\sigma}^2\right)}{g\left(\sigma^{2,(-1)}\right)}\right], \quad \text{where } g(\sigma^2) = \frac{\sqrt{(1-w)^2 + w^2\sigma^2}}{1-w+w\sigma}.$$

The remaining full conditionals used in the Gibbs sampler are all standard. That is,

$$\begin{aligned} (\beta_0, \boldsymbol{\beta}')' \mid \alpha, \sigma^2, \{\lambda_i\}_{i\in\mathcal{C}_0} &\sim N\left(\boldsymbol{b}_1, \boldsymbol{B}_1\right), \\ \lambda_i \mid \alpha, \beta_0, \boldsymbol{\beta}, \sigma^2 &\sim TN_{(0,\infty)}\left(-p_i + \alpha y_i + \beta_0 + \boldsymbol{x}_i'\boldsymbol{\beta}, \sigma^2\right), \end{aligned}$$

where $\boldsymbol{z}_i' = (1, \boldsymbol{x}_i')$, $\boldsymbol{B}_1^{-1} = \sigma^{-2}\sum_{i=1}^n \boldsymbol{z}_i\boldsymbol{z}_i'$, and $\boldsymbol{b}_1 = \sigma^{-2}\boldsymbol{B}_1\sum_{i=1}^n e_i\boldsymbol{z}_i$ (see Equation (7) for e_i).

D Analysis of other parameters

The regression coefficients except for the intercept are summarized in Figure 6. The broader credible interval for α is attributed to the larger number of prefectures of no production. The intercept and standard error of regression are summarized in Figure 7.

References

Abadir, K. M. and J. R. Magnus (2005). *Matrix Algebra.* Econometric exercises 1. Cambridge: Cambridge University Press.

Amemiya, T. (1973). Regression analysis when the dependent variable is truncated normal. *Econometrica 41*(6), 997–1016.

Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors: simple strategies for empirical practice. *Journal of Business & Economic Statistics 19*(1), 2–16.

Bernard, A. B., S. J. Redding, and P. K. Schott. (2010). Multiple-product firms and product switching. *American Economic Review 100*(1), 70–97.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge: Cambridge University Press.

Chib, S. (1992, 79-99). Bayes inference in the Tobit censored regression model. *Journal of Econometrics 51*(1-2).

Chib, S. (2004). Markov chain Monte Carlo technology. In J. E. Gentle, W. Härdle, and Y. Mori (Eds.), *Handbook of Computational Statistics: Concepts and Methods*, pp. 71–102. Berlin; London: Springer.

Cooley, T. F. and S. J. DeCanio (1977). Rational expectations in American agriculture, 1867-1914. *The Review of Economics and Statistics 59*(1), 9–17.

Deschênes, O. and M. Greenstone (2007). The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *American Economic Review 97*(1), 354–385.

Eaton, M. L. and W. D. Sudderth (2010). Invariance of posterior distribution under reparametrization. *Sankhyā 72*, 101–118.

Gamerman, D. and H. F. Lopes (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference* (2nd ed.). Boca Raton: Chapman & Hall/CRC.

Goodwin, T. H. and S. M. Sheffrin (1982). Testing the rational expectations hypothesis in an agricultural market. *The Review of Economics and Statistics 64*(4), 658–667.

Hall, P. and A. Yatchew (2007). Nonparametric estimation when data on derivatives are available. *The Annals of Statistics 35*(1), 300–323.

Hall, R. E. (1988). The relation between price and marginal cost in U.S. industry. *Journal of Political Economy 96*(5), 921–947.

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 186*(1007), 453–461.

Just, R. E., D. Zilberman, and E. Hochman (1983). Estimation of multicrop production functions. *American Journal of Agricultural Economics 65*(4), 770–780.

Kass, R. E. and L. Wasserman (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association 91*(435), 1343–1370.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory.* New York: Oxford University Press.

Mendelsohn, R., W. D. Nordhaus, and D. Shaw (1994). The impact of global warming on agriculture: a Ricardian analysis. *The American Economic Review 84*(4), 753–771.

Moschini, G. (2001). Production risk and the estimation of ex-ante cost functions. *Journal of Econometrics 100*(2), 357–380.

Pfouts, R. W. (1961). The theory of cost and production in the multi-product firm. *Econometrica 29*(4), 650–658.

Pope, R. D. and J.-P. Chavas (1994). Cost functions under production uncertainty. *American Journal of Agricultural Economics 76*(2), 196– 204.

Pope, R. D. and R. E. Just (1996). Empirical implementation of ex ante cost functions. *Journal of Econometrics 72*(1-2), 231–249.

Pörtner, H.-O., D. Roberts, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, and B. Rama (Eds.) (2022). *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge: Cambridge University Press.

Sandmo, A. (1971). On the theory of the competitive firm under price uncertainty. *The American Economic Review 61*(1), 65–73.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. Van Der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 64*(4), 583–639.

Tierney, L. and J. B. Kadane (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association 81*(393), 82–86.

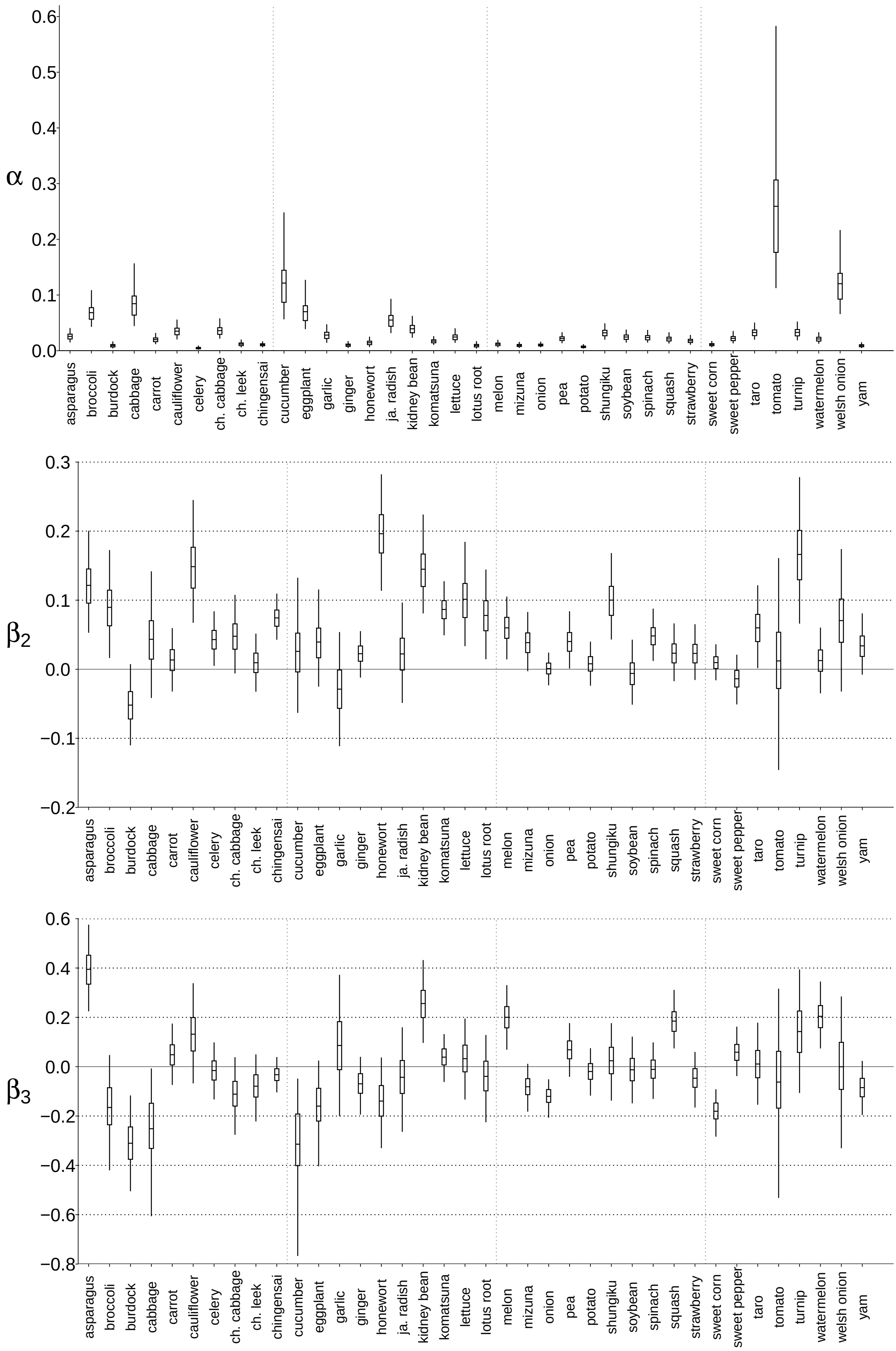

Figure 6: Posterior summary I.

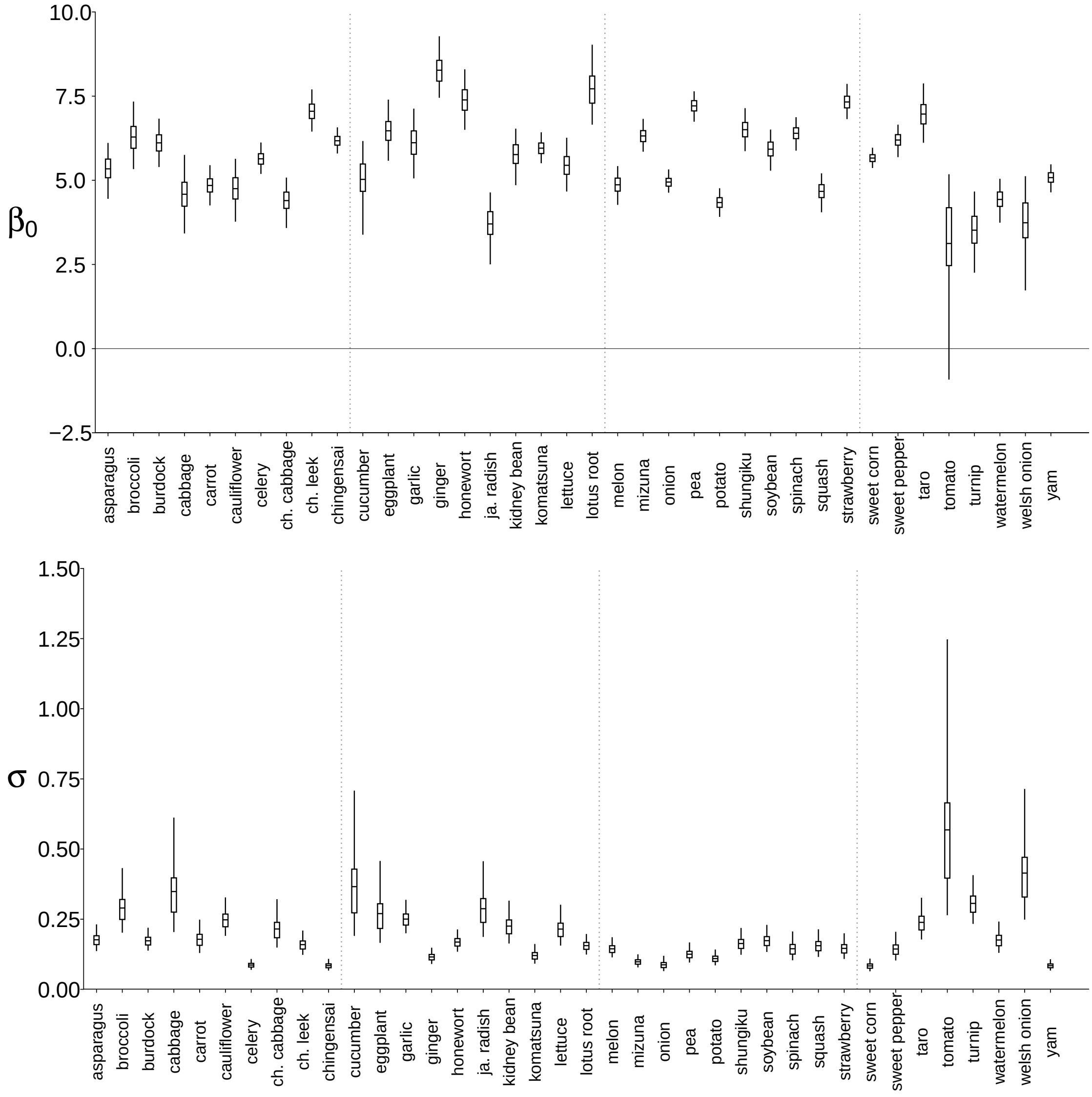

Figure 7: Posterior summary II.