CIRJE-F-1003

Optimal Room Charge and Expected Sales under Discrete Choice Models with Limited Capacity

Taiga Saito
Graduate School of Economics, The University of Tokyo

Akihiko Takahashi
The University of Tokyo

Hiroshi Tsuda
Doshisha University,

March 2016; revised in June 2016

Optimal Room Charge and Expected Sales under Discrete Choice Models with Limited Capacity [*†]

Taiga Saito [‡], Akihiko Takahashi [§], Hiroshi Tsuda [¶]

June 20, 2016

Abstract

In this paper, we introduce a model that incorporates features of the fully transparent hotel booking systems and enables estimates of hotel choice probabilities in a group based on the room charges. Firstly, we extract necessary information for the estimation from big data of online booking for major four hotels near Kyoto station. [1] Then, we consider a nested logit model as well as a multinomial logit model for the choice behavior of the customers, where the number of rooms available for booking for each hotel are possibly limited. In addition, we apply the model to an optimal room charge problem for a hotel that aims to maximize its expected sales of a certain room type in the transparent online booking systems. We show numerical examples of the maximization problem using the data of the four hotels of November 2012 which is a high season in Kyoto city. This model is useful in that hotel managers as well as hotel investors, such as hotel REITs and hotel funds, are able to predict the potential sales increase of hotels from online booking data and make use of the result as a tool for investment decisions.

Keywords: Hotels in Kyoto, Revenue management, Online booking, Discrete choice model

1 Introduction

Hotel revenue management is an important issue not only for hotel managers, but also for hotel REITs and tourism funds who invest in hotels and engage in the management in order to increase profitability from the properties. By the revenue management, they

[*]Forthcoming in *International Journal of Hospitality Management.*

[†]All the contents expressed in this research are solely those of the authors and do not represent any views or opinions of any institutions. The authors are not responsible or liable in any manner for any losses and/or damages caused by the use of any contents in this research.

[‡]Financial Research Center at Financial Services Agency, Government of Japan. 3-2-1 Kasumigaseki, Chiyoda-ku, Tokyo 100-8967, Japan. saitotaiga@hotmail.com

[§]Graduate School of Economics, The University of Tokyo.

[¶]Department of Mathematical Sciences, Doshisha University, Visiting Professor at the Institute of Statistical Mathematics.

[1]The data were provided by National Institute of Informatics.

are able to know how they operate the hotels in order to increase the revenues. For the investors in hotels, by knowing the sizes of potential sales increase, the revenue management can also be used as an investment decision making tool. Taking this into account, we consider maximization of expected sales of a certain room type of hotels in a group in the same area by solving for optimal room charges. The model incorporates random choice of hotels by customers who visit an online hotel booking website. Then we use big data of online booking for major four hotels near Kyoto station for the estimation, which were collected from a Japanese booking website by National Institute of Informatics. They include hotel names, accommodation plans with room types, remaining numbers of the plans available for booking, and prices of the plans for fourteen days booking periods prior to check-in dates. They are meaningful since those include more detailed information than the financial statements which are usually undisclosed. We first extract necessary information from the data under some suitable assumptions so that it can be used to estimate the parameters of the model.

For literatures on online hotel booking systems, Wang et al. [27] investigates relations between the quality of hotel websites and customers' online booking intentions. Noone and Mattila [25] studies how the two ways of rate presentation for multiple-day stays, the blended and non-blended presentations of best available rates, affect customers' willingness to book. Liu and Zhang [16] examines factors for travelers to choose an online booking channel among hotel and online travel agencies websites. Casaló et al. [7] studies effects of online hotel ratings by online travel communities, such as TripAdvisor, on booking behaviors of customers. Ladhari and Michaud [15] investigates influences of online word of mouth in social network services, such as Facebook, on the choice of a hotel. Abrate et al. [1] examines dynamic pricing strategies of 1,000 European hotels. Viglia et al. [29] analyses the impact of hotel price sequences on customers' reference price which is used to evaluate market prices.

For applications of discrete choice models to tourism research, Masiero et al. [17], [20], Nicolau and Masiero [24], and Masiero and Nicolau [18] use the mixed logit model for the analysis of choice behaviors of hotel customers. Masiero et al. [17] investigates customers' willingness to pay for hotel room attributes within a single hotel property. Masiero et al. [20] examines an asymmetric preference for hotel room choices based on prospect theory. Masiero and Nicolau [18] analyses the determinants of individual price sensitivities to tourism activities. Nicolau and Masiero [24] observes the effect of individual price sensitivities to tourism activities on on-site expenditures. Moreover, Viglia et al. [28] investigates impact of customer reviews on preference for hotel choice by a rank conjoint experiment. Masiero et al. [19] examines determinant factors of tourist expenditure for accommodation by a quantile regression.

Also, there are some other literatures on hotel revenue management by quantitative approaches. Quantitative revenue management has been studied mostly on online booking systems of airline tickets. For instance, Kimes [13], Weatherford and Bodily [26] summarize the methodologies. For hotels, Bitran and Mondschein [6], Badinelli [4] consider sales maximization of a single hotel. In particular, Bitran and Mondschein [6] takes into consideration the case of multi-day stays, and Badinelli [4] investigates dynamic room pricing where the optimal policy depends on the vacancy and the remaining days before the check-in date. Anderson and Xie [3] is the first study that deals with the sales maximization of hotels in a group, taking into account choice behaviors of customers in online booking systems. Specifically, Anderson and Xie [3] studies on opaque booking systems where the name of a hotel that customers try to book is concealed until the booking is done, assuming the nested-logit model for the customer choices among hotels in different areas and price ranges. For qualitative analysis on practices of hotel revenue management, see Baker and Collier [5], Donaghy et al. [8], Hanks et al. [10], Kims [14], Lieberman [23].

While Anderson and Xie [3] covers the opaque booking systems which are popular in the United States, the fully transparent booking systems, where customers consider booking by knowing names of the hotels, are common in Japan. This is because customers care for names and reviews of the hotels, since the quality of the services are dissimilar even among the same rank of hotels and some hotels differentiate themselves from the rivals by selling variety of accommodation plans which include options such as tickets for sightseeing facilities, recreation, and meals.

In the transparent booking systems, particularly in high seasons, as customers are only able to choose hotels offering available rooms, the limitation of the number of rooms that the rival hotels can offer is an important factor to be considered in modeling. On the other hand, in the opaque booking systems, choice categories, a pair of ratings and areas, are not exhausted as long as some of the hotels in the categories provide rooms. Hence we model the transparent booking systems, taking into account the limitation of the available number of rooms for the rival hotels. In the model, fully-occupied hotels are excluded from the choice alternatives. We note that our model contains so called a waterfall structure as in collateralized debt obligations in finance. (See Gibson [9], Hull and White [12] for details.)

Moreover, in contrast to the opaque booking systems where the daily changing pricing is a key to maximizing the sales, consistent pricing is important in the transparent booking systems. As Anderson and Xie [3] points out, the frequent price changes and discounts in a booking period in the fully transparent booking systems may lose loyal customers of the hotel who book in advance thorough the direct selling channel at regular high prices. Therefore, our work aims to obtain one optimal room charge, which is unchanged through a booking period in the transparent booking systems, while Anderson and Xie [3] deals with daily pricing in the opaque booking systems in the United States where hotels maximize their profits by selling out their rooms by discounting in several days before the check-in date.

The organization of the paper is as follows. Section 2 introduces the model that reflects choice behavior of customers in the transparent online booking systems. The model assumes a Poisson process for the frequency of visiting of customers and a nested logit model, as well as a multinomial logit model, with limited number of available rooms for hotels. An algorithm to calculate the expected sales under the model is also shown in this section. Section 3 provides numerical examples of the optimal room charge. Finally, Section 4 concludes. Appendices provide properties on the nested logit model in Section 2.1, and proofs of the theorem on existence of an optimal room charge in Section 2.2 and the lemma in Section 3.2.

2 Model and Estimation

2.1 Model Specification

In this subsection, we introduce the model that describe booking activities of customers in an online system who make a reservation for a certain type of room in a group of hotels. Let $\{1, 2, \ldots, M\}$ be check-in dates. First fix a check-in date $m,\ 1 \leq m \leq M$. Let $[0, T]$ be a booking period for the check-in date, where 0 is the start date of the period and T is a check-in date. We fix the check-in date. Let $t \in [0, T]$ be a booking date. We assume that the total number of rooms booked during the period for the group for the check-in date follows a Poisson process $\{N_t^{(m)}\}_{0 \leq t \leq T}$ with intensity $\lambda^{(m)}$. We further assume the following.

- Fix the number of hotels in the group. Let $L \in \mathbf{N}$ be the number of hotels and we name the hotels from hotel 1 to hotel L.

- Let $R_t^i, (1 \leq i \leq L)$ be the number of rooms of hotel i booked until time t that satisfies

$$\sum_{i=1}^{L} R_t^i = N_t^{(m)}, \tag{1}$$

$$0 \leq R_t^i \leq q_i^{(m)}. \tag{2}$$

  Here $q_i^{(m)} \in \mathbf{N} \cup \{\infty\}$ is the maximum number of rooms available for booking for hotel i $(0 \leq i \leq L)$.

- Let $\gamma \in \Gamma := \Pi_{i=1}^{L}\{0, 1\}$ be the state of full occupancy. That is, 0 for the i-th component of γ means that there is no available room for the hotel i, otherwise the hotel is available for booking.

- Let $p_i^{(m)\gamma}$ $(0 \leq p_i^{(m)\gamma} \leq 1,\ 0 \leq i \leq L)$ be the choice probability of hotel i by a customer under the occupancy state γ, which satisfies

$$\sum_{i=1}^{L} p_i^{(m)\gamma} = 1. \tag{3}$$

- Let $x_i^{(m)}$ be the room charge of hotel i for the check-in date m and $V_i^{(m)} : \mathbf{R} \to \mathbf{R},\ (i = 1, 2, \ldots, L)$ be linear functions of $x_i^{(m)}$ defined by

$$V_i^{(m)} = \alpha_i + \beta_i x_i^{(m)} + \delta_i y^{(m)},\ (i = 1, \ldots, L), \tag{4}$$

$$\beta_i < 0. \tag{5}$$

  where $y^{(m)}$ is a dummy variable that takes 1 if the check-in date m is a day before a holiday or 0 otherwise.

We note that in the stochastic utility maximization theory as in McFadden [21], $V_i^{(m)}$ corresponds to the deterministic term of the random utility $U_i^{(m)}$ of a customer for choosing hotel i with decomposition $U_i^{(m)} = V_i^{(m)} + \epsilon_i^{(m)}$, where $\epsilon_i^{(m)}$ is the random term of the utility. In the theory, since a customer choose hotel i with the highest utility against alternatives, the choice probability for hotel i which depends on the distribution of $(\epsilon_1^{(m)}, \ldots, \epsilon_L^{(m)})$ is given as

$$\mathbf{P}(U_i^{(m)} > U_j^{(m)},\ i \neq j). \tag{6}$$

In particular, Theorem 1 in McFadden [21] provides so called *the generalized extreme value model*, a class of probabilistic choice models, which is consistent with utility maximization.

Hereafter, as choice probabilities we take a nested logit model (e.g. Section 9.3.5 in Amemiya [2]), which is originally developed by McFadden [21] as an example of generalized extreme value models. (e.g. Section 5 in McFadden [21], Section 5.15 in McFadden [22])

Assumption 1. *(Choice probability) Let $C_1, \ldots, C_n$ be disjoint subsets of $\{1, \ldots, L\}$ satisfying*

$$\bigcup_{k=1}^{n} C_k = \{1, \ldots, L\},\ C_k \cap C_l = \emptyset. \tag{7}$$

Let $k_i \in \{1, \ldots, n\}$ be the index such that $i \in C_{k_i}$ and $\{j \in C_{k_i} | \gamma_j = 1\} \neq \emptyset$. The choice probability of hotel i under the occupancy state $\gamma \in \Gamma$ is given by

$$p_i^{(m)\gamma} = \frac{\{\sum_{j \in C_{k_i}} \exp(\frac{V_j^{(m)}}{\nu_{k_i}}) 1_{\{\gamma_j = 1\}}\}^{\nu_{k_i}}}{\sum_{k=1}^{n} \{\sum_{j \in C_k} \exp(\frac{V_j^{(m)}}{\nu_k}) 1_{\{\gamma_j = 1\}}\}^{\nu_k}} \cdot \frac{\exp(\frac{V_i^{(m)}}{\nu_{k_i}}) 1_{\{\gamma_i = 1\}}}{\sum_{j \in C_{k_i}} \exp(\frac{V_j^{(m)}}{\nu_{k_i}}) 1_{\{\gamma_j = 1\}}} \quad (i = 1, \ldots, L), \tag{8}$$

where $\nu_1, \ldots, \nu_n$ are constants taking values $0 < \nu_1, \ldots, \nu_n \leq 1$. We set $p_i^{(m)\gamma} = 0$ when $\{j \in C_{k_i} | \gamma_j = 1\} = \emptyset$.

We notice that these choice probabilities correspond to the following joint distribution function of the random terms of utilities, $(\epsilon_1^{(m)}, \ldots, \epsilon_L^{(m)})$ (e.g. Section 5 of McFadden [21]): for all $w_1, \ldots, w_L \in \mathbf{R}$,

$$F(w_1, \ldots, w_L) = \exp\left(-G\left(\exp(-w_1), \ldots, \exp(-w_L)\right)\right), \tag{9}$$

$$G(z_1, \ldots, z_L) = \sum_{k=1}^{n} (\sum_{j \in C_k} z_j^{-\nu_k} 1_{\{\gamma_j = 1\}})^{\nu_k}. \tag{10}$$

Particularly, when $\nu_1 = \cdots = \nu_n = 1$, the choice probabilities agree with the ones in a multinomial logit model:

$$p_i^{(m)\gamma} = \frac{\exp(V_i^{(m)}) 1_{\{\gamma_i = 1\}}}{\sum_{j=1}^{L} \exp(V_j^{(m)}) 1_{\{\gamma_j = 1\}}},\ (i = 1, \ldots, L). \tag{11}$$

Remark 1. *The parameters $1 - \nu_k$ can be interpreted as the degrees of the similarity within a set C_k. (For instance, see Section 5 in McFadden [21] and Section 5.15 in McFadden [22] for the details.): when $\nu_k = 1$, the choices in the set are dissimilar,* *and as ν_k decreases to 0, the choices in C_k become similar. More precisely, as in well-known properties of a generalized mean (or a power mean), when ν_k goes to 0, the term $\left\{\sum_{j\in C_k}\exp(\frac{V_j^{(m)}}{\nu_k})1_{\{\gamma_j=1\}}\right\}^{\nu_k}$ in (8) is nondecreasing on $0<\nu_k\le 1$, and converges to $\max_{j\in\tilde{C}_k}\exp(V_j^{(m)})$:*

$$\lim_{\nu_k\to 0}\left\{\sum_{j\in C_k}\exp(\frac{V_j^{(m)}}{\nu_k})1_{\{\gamma_j=1\}}\right\}^{\nu_k}=\max_{j\in\tilde{C}_k}\exp(V_j^{(m)}),$$

where $\tilde{C}_k=\{j\in C_k|\gamma_j=1\}\neq\emptyset$. For convenience, we give the proofs of those properties in Appendix A.

2.2 Expected Sales of Hotel i

In this subsection, we observe how the expected sales of hotel i is calculated. Let $S_k:=\Pi_{j=1}^k\{1,2,\dots,L\}$ be a set of booking orders when there are k bookings in total until the check-in date. For each scenario $(i_1,\dots,i_k)\in S_k$, the corresponding scenario of the occupancy state $(\gamma^{(m)1}_{(i_1,\dots,i_k)},\gamma^{(m)2}_{(i_1,\dots,i_k)},\dots,\gamma^{(m)k}_{(i_1,\dots,i_k)})\in\Pi_{j=1}^k\Gamma$ is defined as follows.

$$\gamma^{(m)j}_{(i_1,\dots,i_k)}:=(\gamma^{(m)j}_{1,(i_1,\dots,i_k)},\dots,\gamma^{(m)j}_{L,(i_1,\dots,i_k)}),\ (j=1,\dots,k) \tag{12}$$

$$\gamma^{(m)j}_{i,(i_1,\dots,i_k)}:=\begin{cases}1, \text{ if } \sum_{l=1}^{j-1}1_{\{i_l=i\}}<q_i^{(m)},\\ 0, \text{ otherwise.}\end{cases} \tag{13}$$

Hereafter, for the notational simplicity, we abbreviate $\gamma^{(m)j}_{(i_1,\dots,i_k)}$ and $\gamma^{(m)j}_{i,(i_1,\dots,i_k)}$ as $\gamma^{j}_{(i_1,\dots,i_k)}$ and $\gamma^{j}_{i,(i_1,\dots,i_k)}$, respectively. We assume that the conditional probability of the occurrence of the scenario $(i_1,\dots,i_k)\in S_k$ under $N_T^{(m)}=k$ is given by

$$\mathbf{P}(\{(i_1,\dots,i_k)\}|N_T^{(m)}=k)=p_{i_1}^{(m)\gamma^1_{(i_1,\dots,i_k)}}\dots p_{i_k}^{(m)\gamma^k_{(i_1,\dots,i_k)}}. \tag{14}$$

This has the following interpretations. For check-in date m, when the total number of booking until the check-in date determined by the Poisson random variable $N_T^{(m)}$ is k, the choice probability $p_{i_l}^{(m)\gamma^l_{(i_1,\dots,i_k)}}$ at the l-th booking $(1\le l\le k)$ only depends on the occupancy state $\gamma^l_{(i_1,\dots,i_k)}$. Given the occupancy state, customers choose a hotel irrespectively of the choices previously made. We notice that this mechanism where the number of alternatives decreases as they are filled up is analogous to a waterfall structure, a rule of prioritized cash payments in collateralized debt obligations. (See Gibson [9], Hull and White [12] for details.)

Then, the conditional expectation of the number of rooms booked for hotel i under $N_T^{(m)}=k$ is

$$\mathbf{E}[R_T^i|N_T^{(m)}=k]=\sum_{(i_1,\dots,i_k)\in S_k}a^i_{(i_1,\dots,i_k)}p_{i_1}^{(m)\gamma^1_{(i_1,\dots,i_k)}}\dots p_{i_k}^{(m)\gamma^k_{(i_1,\dots,i_k)}}, \tag{15}$$

where

$$a^i_{(i_1,\ldots,i_k)} = \sum_{l=1}^{k} 1_{\{i_l=i\}}. \tag{16}$$

Therefore, the expectation of the sales of hotel i at the check-in date T is calculated as follows.

$$\begin{aligned}
\mathbf{E}[x_i^{(m)} R_T^i] &= x_i^{(m)} \mathbf{E}[R_T^i] \\
&= x_i^{(m)} \sum_{k=0}^{\infty} \mathbf{E}[R_T^i | N_T^{(m)} = k] \mathbf{P}(N_T^{(m)} = k) \\
&= x_i^{(m)} \sum_{k=0}^{\infty} \mathbf{E}[R_T^i | N_T^{(m)} = k] \exp(-\lambda^{(m)} T) \frac{(\lambda^{(m)} T)^k}{k!} \\
&= x_i^{(m)} \sum_{k=0}^{\infty} \sum_{(i_1,\ldots,i_k)\in S_k} a^i_{(i_1,\ldots,i_k)} p_{i_1}^{(m)\gamma^1_{(i_1,\ldots,i_k)}} \cdots p_{i_k}^{(m)\gamma^k_{(i_1,\ldots,i_k)}} \exp(-\lambda^{(m)} T) \frac{(\lambda^{(m)} T)^k}{k!}.
\end{aligned} \tag{17}$$

In the nested logit model, the expected sales attains its maximum at some $x_i^{(m)} \in [0, \infty)$ as in the following theorem. We provide the proof in Appendix B.

Theorem 1. *Suppose that there exists $j \neq i$ such that $q_j^{(m)} = \infty$. Then, in the nested logit model (8), the expected sales (17) as a function of $x_i^{(m)}$ has its maximum point in $[0, \infty)$.*

Note that this assumption that there exists $j \neq i$ such that $q_j^{(m)} = \infty$ indicates that there is at least one hotel other than hotel i, which has enough capacity and can never be fully-occupied. In the case where there is no such hotel, if all the other hotels are fully-booked, the choice probability of hotel i becomes 1 regardless of the room charge of hotel i. In such a case, the expected sales is increasing with respect to $x_i^{(m)}$ and does not have a maximum point.

3 Numerical Example

In this section, we show numerical examples of the maximization of the expected sales by the multinomial logit model and the nested logit model with limited capacity introduced in Section 2. We use empirical online booking data of major four hotels near Kyoto station which were offered by National Institute of Informatics and examine the maximization problem of the four hotels. We consider choice behaviors of customers who seek to stay at a hotel very near Kyoto station, where we assume that they choose a hotel among the major four hotels close to the station. In Kyoto, the station neighborhood is a major accommodation area and the four hotels are luxury full-service hotels. Hence, it is reasonable to consider that there exists a group of customers who prefer to choose a decent high-class full-service hotel among the four in this area. The validity of selecting the four hotels for the choices in the analysis will be discussed in more detail in Section

3.3. For the room type customers choose in the model, we adopt nonsmoking standard twin room, a common room type for tourists. We first estimate parameters of the models with limited capacity by the maximum likelihood method on the data of November in 2012, which is the busiest season in Kyoto when tourists enjoy watching the fall foliage. Especially, since the latter half of November is the peak season of the fall foliage and the hotels set high room charges regardless of whether a check-in date is a day before a holiday, we treat the last two weeks in November 2012 as a day before a holiday. We also note that the Japanese yen (JPY) had been strong against other currencies in 2012, and the year was right before the number of tourists from overseas started to rise and the hotels were constantly occupied by these tourists. Hence, in order to exclude the large effect of overseas travelers on booking, we use the data of this year. Then we consider maximization of the expected sales of the hotels for two check-in dates, a normal weekday and a day before a holiday including the last half of November.

In the following of this section, after Section 3.1 explains details of the data set used in the numerical example, Section 3.2 provides estimation results of the parameters in the multinomial logit model and the nested logit model, and Section 3.3 discusses validity of the data set in the analysis. Finally, Section 3.4 shows the optimal room charges and the expected sales for the two types of check-in dates calculated with the estimated nested logit model.

3.1 Data Set

We use online booking data of major four hotels near Kyoto station, which were collected from a Japanese booking website by National Institute of Informatics. The four hotels are named as Hotel A, Hotel B, Hotel C and Hotel D in this paper. The original data include available numbers for booking for all accommodation plans and their prices for each booking date of all check-in dates.

Note that the original data contain no information on booked numbers either per room type or per accommodation plan. Moreover, as the available number for booking of an accommodation plan changes in tandem with other plans that offer the same room type, we do not know which plan was sold or how many rooms were booked for the room type when the available number for booking decreases. Thus, we have to estimate the booked number on a specific room type in some way. In addition, since a room type is overlapped among various accommodation plans and their prices are different, in order to comply with the model, we have to define a representative price of the room type for a check-in date.

Considering these points, we initially convert the original data to a data set used for our analysis in the following way. For the number of rooms booked for a specific room type, we first calculate changes from the previous day in the available number of booking for each accommodation plan. If the change is a decrease, we assume that this number of plans were booked. If the available number of booking is unchanged, we assume that no plan was booked, and if the change is an increase, this number of plans were just supplied by the hotel. Then we take a maximum of these numbers among all accommodation plans that offer this room type. We assume the maximum as the number of rooms booked for this room type.

Next, since in practice, hotels occasionally change their room charges in some degree during a booking period, and there do not exist single prices as in the model for the check-in dates, we take an average of the prices over the booking period for each accommodation plan. Then, we take a minimum of the average prices among all the accommodation plans that offer this room type, that is, nonsmoking standard twin room, in order to find the average price with no or minimal options. We assume the minimum as the representative room charge over the booking period for the room type. This will be denoted by $x_i^{(m)}$ in the following.

The data set includes booking information of the 30 check-in days in November 2012 for the standard nonsmoking twin room type of the four hotels. For each check-in date, the numbers of rooms booked for 14 booking dates ranging from the check-in date to 13 days prior to the check-in date, and the representative room charges over the booking period are available. Table 1 shows an example of the booking data for the check-in date of a weekday in November 2012. The right side of the table shows the numbers of rooms sold for this room type for each hotel on each booking date. "Full" in Table 1 indicates that no room was available for booking as the rooms were sold out before the date.

	Hotel A	Hotel B	Hotel C	Hotel D
check-in date	3	Full	0	5
-1 day	0	Full	1	2
-2 days	2	Full	0	0
-3 days	1	Full	11	2
-4 days	0	3	0	0
-5 days	0	0	0	0
-6 days	0	2	0	1
-7 days	0	0	0	0
-8 days	1	2	0	4
-9 days	1	4	0	5
-10 days	1	0	18	1
-11 days	3	0	2	0
-12 days	0	2	0	8
-13 days	0	0	0	0
number of rooms booked	12	13	32	28

Table 1: Estimated Numbers of Rooms Booked for Check-In Date, a weekday in November 2012 with Representative Room Charges 11,985, 11,000, 19,938, 18,000 in JPY for Hotel A, Hotel B, Hotel C, and Hotel D

Table 2 summarizes the estimated data set. Out of a total of 1,680 (= 30days $\times$ 14days $\times$ 4hotels) booking days for the four hotels, 1,273 days were available for booking, and the rest of the days were unavailable due to the full occupancy. In the available booking days, 1,856 rooms were booked in total. Note that the room charges in Table 2 and prices hereafter are all expressed in JPY.

	Hotel A	Hotel B	Hotel C	Hotel D
maximum representative room charge per check-in date	24,252	18,500	32,329	30,030
minimum representative room charge per check-in date	11,292	9,571	18,246	15,286
average room charge for whole check-in dates	17,502	15,071	27,246	21,170
maximum representative room charge per check-in date, a weekday	16,886	18,500	27,293	22,286
minimum representative room charge per check-in date, a weekday	11,292	9,571	18,246	15,286
average representative room charge per check-in date, a weekday	13,858	13,149	22,020	19,612
maximum representative room charge per check-in date, a day before a holiday	24,252	18,085	32,329	30,030
minimum representative room charge per check-in date, a day before a holiday	15,700	15,236	26,277	18,857
average representative room charge per check-in date, a day before a holiday	21,175	17,321	30,387	22,122
maximum number of rooms booked per check-in date	61	15	37	72
minimum number of rooms booked per check-in date	0	0	0	0
average number of rooms booked for per check-in date	15.33	5.03	10.27	31.23
total number of rooms booked for whole check-in dates	460	151	308	937
maximum number of available booking days per check-in date	14	14	14	14
minimum number of available booking days per check-in date	0	0	0	12
average number of available booking days per check-in date	9.93	8.07	10.57	13.87
total number of available booking days for whole check-in dates	298	242	317	416
maximum number of accommodation plans per check-in date	36	11	5	38
minimum number of accommodation plans per check-in date	32	11	5	2
average number of accommodation plans per check-in date	34.14	11.00	5.00	20.93

Table 2: Estimated Data Summary, November 2012.

3.2 Estimation

In this subsection, we provide estimation results of the multinomial logit model and the nested logit model with the data described in the previous subsection.

First, the next lemma indicates that in the nested logit model as well as the multinomial logit model, the choice probability is reexpressed as follows, whose proof is given in Appendix C.

Lemma 1. *Let*

$$\tilde{V}_1^{(m)} = \beta_1 x_1^{(m)}, \tag{18}$$

$$\tilde{V}_i^{(m)} = \tilde{\alpha}_i + \beta_i x_i^{(m)} + \tilde{\delta}_i y^{(m)} \ (i = 2, \ldots, L), \tag{19}$$

where $\tilde{\alpha}_i = \alpha_i - \alpha_1, \tilde{\delta}_i = \delta_i - \delta_1$.

Then, $p_i^{(m)\gamma}$ *in the nested logit model (8) is rewritten as follows.*

$$p_i^{(m)\gamma} = \frac{\{\sum_{j\in C_{k_i}} \exp(\frac{\tilde{V}_j^{(m)}}{\nu_{k_i}})1_{\{\gamma_j=1\}}\}^{\nu_{k_i}}}{\sum_{k=1}^{n}\{\sum_{j\in C_k} \exp(\frac{\tilde{V}_j^{(m)}}{\nu_k})1_{\{\gamma_j=1\}}\}^{\nu_k}} \cdot \frac{\exp(\frac{\tilde{V}_i^{(m)}}{\nu_{k_i}})1_{\{\gamma_i=1\}}}{\sum_{j\in C_{k_i}} \exp(\frac{\tilde{V}_j^{(m)}}{\nu_{k_i}})1_{\{\gamma_j=1\}}}, \ (i = 1, \ldots, L). \tag{20}$$

In particular, when $\nu_{k_1}, \ldots, \nu_{k_L} = 1$ *(the multinomial logit model (11)),*

$$p_i^{(m)\gamma} = \frac{\exp(\tilde{V}_i^{(m)})1_{\{\gamma_i=1\}}}{\sum_{j=1}^{L} \exp(\tilde{V}_j^{(m)})1_{\{\gamma_i=1\}}} \ (i = 1, \ldots, L). \tag{21}$$

Then, Table 3,4,5,6 show the estimation results of the coefficients of the explanatory variables in the nested logit model (20) and the multinomial logit model (21).

	Estimate	Std. Error	t-value	p-value
$\tilde{\alpha}_2$, Hotel B:(intercept)	-0.7151	0.5097	-1.4029	0.161
$\tilde{\alpha}_3$, Hotel C:(intercept)	0.1455	0.5506	0.2642	0.792
$\tilde{\alpha}_4$, Hotel D:(intercept)	0.8201	0.3450	2.3771	0.017
$\tilde{\delta}_2$, Hotel B:holiday	-1.0346	0.3104	-3.3331	0.001
$\tilde{\delta}_3$, Hotel C:holiday	-0.6338	0.3169	-2.0003	0.045
$\tilde{\delta}_4$, Hotel D:holiday	-0.5633	0.2055	-2.7416	0.006
β_1, Hotel A:price	-0.000125	0.0000330	-3.7733	0.000
β_2, Hotel B:price	-0.000122	0.0000457	-2.669	0.008
β_3, Hotel C:price	-0.000099	0.0000267	-3.6976	0.000
β_4, Hotel D:price	-0.000118	0.0000268	-4.3953	0.000
ν_1, iv.Hotel A,Hotel D	0.670	0.179	3.7473	0.000
ν_2, iv.Hotel B,Hotel C	0.594	0.189	3.1389	0.002

Table 3: Estimation Results, Nested Logit Model, November 2012.

Log-Likelihood	-1956.2
McFadden R^2	0.1165
Likelihood ratio test	$\chi^2 = 515.91$ (p.value $\leq$ 2.22e-16)

Table 4: Estimation Results (Tests), Nested Logit Model, November 2012.

	Estimate	Std. Error	t-value	p-value
$\tilde{\alpha}_2$, Hotel B:(intercept)	-0.5539	0.7207	-0.7686	0.442
$\tilde{\alpha}_3$, Hotel C:(intercept)	0.4013	0.6028	0.6658	0.506
$\tilde{\alpha}_4$, Hotel D:(intercept)	1.1090	0.4651	2.3846	0.017
$\tilde{\delta}_2$, Hotel B:holiday	-1.1303	0.4127	-2.7388	0.006
$\tilde{\delta}_3$, Hotel C:holiday	-0.6523	0.2956	-2.2064	0.027
$\tilde{\delta}_4$, Hotel D:holiday	-0.7378	0.2404	-3.0692	0.002
β_1, Hotel A:price	-0.000167	0.0000307	-5.4509	0.000
β_2, Hotel B:price	-0.000185	0.0000570	-3.2465	0.001
β_3, Hotel C:price	-0.000129	0.0000272	-4.7452	0.000
β_4, Hotel D:price	-0.000157	0.0000214	-7.3216	0.000

Table 5: Estimation Results, Multinomial Logit Model, November 2012.

Log-Likelihood	-1959
McFadden R^2	0.11522
Likelihood ratio test	$\chi^2 = 510.22$, (p.value $\leq$ 2.22e-16)

Table 6: Estimation Results (Tests), Multinomial Logit Model, November 2012.

We estimate the parameters in the models by the maximum likelihood method with the maximum likelihood functions of the following forms:

(the nested logit model)

$$\Pi_{m=1}^{30}\Pi_{l=0}^{13}\Pi_{i=1}^{4}\left\{p_i^{(m)\gamma_{(m,l)}}(x_1^{(m)},\ldots,x_4^{(m)},y^{(m)};\tilde{\alpha}_2,\tilde{\alpha}_3,\tilde{\alpha}_4,\beta_1,\ldots,\beta_4,\tilde{\delta}_2,\tilde{\delta}_3,\tilde{\delta}_4,\nu_1,\nu_2)\right\}^{\eta_{(m,l,i)}}. \tag{22}$$

(the multinomial logit model)

$$\Pi_{m=1}^{30}\Pi_{l=0}^{13}\Pi_{i=1}^{4}\left\{p_i^{(m)\gamma_{(m,l)}}(x_1^{(m)},\ldots,x_4^{(m)},y^{(m)};\tilde{\alpha}_2,\tilde{\alpha}_3,\tilde{\alpha}_4,\beta_1,\ldots,\beta_4,\tilde{\delta}_2,\tilde{\delta}_3,\tilde{\delta}_4)\right\}^{\eta_{(m,l,i)}}. \tag{23}$$

In this estimation, we assume the following. Let hotel 1, hotel 2, hotel 3, hotel 4 be Hotel A, Hotel B, Hotel C and Hotel D respectively. Let check-in date 1, ..., check-in date 30 be the 30 check-in dates in November 2012 in ascending order. For each check-in date, there are 14 days of the booking period ranging from the check-in date to the 13 days prior to the check-in date. We call the booking date, which is l days prior to the check-in date ($0 \leq l \leq 13$), booking date l. Let $x_1^{(m)}, x_2^{(m)}, x_3^{(m)}, x_4^{(m)}$ ($1 \leq m \leq 30$) be the representative room charge defined in Section 3.1 of the nonsmoking standard twin room of hotel 1, hotel 2, hotel 3, hotel 4 over the booking period for check-in date m. Let $\gamma_{(m,l)}$ be the occupancy state of booking date l for the check-in date m, and $\eta_{(m,l,i)}$ be the number of rooms booked for hotel i on booking date l for check-in date m. Also, $p_i^{(m)\gamma_{(m,l)}}(x_1^{(m)},\ldots,x_4^{(m)},y^{(m)};\tilde{\alpha}_2,\tilde{\alpha}_3,\tilde{\alpha}_4,\beta_1,\ldots,\beta_4,\tilde{\delta}_2,\tilde{\delta}_3,\tilde{\delta}_4,\nu_1,\nu_2)$ and $p_i^{(m)\gamma_{(m,l)}}(x_1^{(m)},\ldots,x_4^{(m)},y^{(m)};\tilde{\alpha}_2,\tilde{\alpha}_3,\tilde{\alpha}_4,\beta_1,\ldots,\beta_4,\tilde{\delta}_2,\tilde{\delta}_3,\tilde{\delta}_4)$ represent the choice probabilities (20) and (21), respectively.

For the nested logit model, we assume that the alternatives are grouped into two nests: Hotel A and Hotel D which offer a large number of accommodation plans in online booking systems, more than 30 plans at a maximum and more than 20 plans on average for the room type in November 2012 as in Table 2 for example, and Hotel B and Hotel C which sell several accommodation plans. Hence we set $C_1 = \{1,4\}, C_2 = \{2,3\}$. We remark that we also estimated the nested logit model with the other possible nesting structures. In any of these cases, the parameters were not estimated properly. The results were either calculation failure of the maximum likelihood method or inappropriate parameters with ν_k greater than 1.

In Table 3 and 5, β_i ($i = 1,2,3,4$) show all negative signs. This agrees with the intuition that as the room charge of a hotel increases, the utility of customers for the hotel decreases. The negative signs of the coefficients of the holiday factor $\tilde{\delta}_i$ ($i = 2,3,4$) indicate that for check-in dates, each of which is a day before a holiday, an increase of the utility for Hotel B,C,D from weekday check-in dates is less than that of Hotel A. Moreover, the low absolute values of t-values for $\tilde{\alpha}_2, \tilde{\alpha}_3$ in Table 3 and 5 indicate that the intercepts α_i are not very different between Hotel A and Hotel B or Hotel A and Hotel C. However, $\tilde{\alpha}_2, \tilde{\alpha}_3$ have a statistical significance when the explanatory variables x_i are centered. For example, if we shift x_i by 16,020, which is the average of all eight minimum representative room charges in Table 2, that is, if we set $\hat{x}_i^{(m)} = x_i^{(m)} - 16{,}020$,

($i = 1, 2, 3, 4$) and estimate the parameters with

$$\hat{V}_1^{(m)} = \hat{\beta}_1 \hat{x}_1^{(m)}, \tag{24}$$

$$\hat{V}_i^{(m)} = \hat{\alpha}_i + \hat{\beta}_i \hat{x}_i^{(m)} + \hat{\delta}_i y^{(m)} \ (i = 2, 3, 4), \tag{25}$$

we observe that the t-values of all the intercepts in the nested logit model as well as the multinomial logit model improve, and the lowest absolute value of a t-value is 1.89 of $\hat{\alpha}_3$ in the nested logit model. This implies that when the prices are centered at the level, the intercepts have a significance.

We examine the intercepts in the centered prices case in Section 3.3 by computing the choice probabilities when all the prices are at the same levels.

Also, as stated in Remark 1 in Section 2.1, $1 - \nu_1$ (~ 0.33) and $1 - \nu_2$ (~ 0.41) can be interpreted as measures of the similarity of the two alternatives in C_1 and C_2, respectively. Moreover, it is known that for the case of two alternatives within the same set C_k, $1 - \nu_k^2$ is equivalent to the correlation between the two alternatives. (e.g. p.300 in Amemiya [2].) Hence, the correlations between two alternatives are around 0.55 and 0.65 within C_1 and C_2, respectively.

Next, we conduct the Hausman test for the multinomial logit model in Hausman and McFadden [11]. The Hausman test judges whether the IIA (Independence from Irrelevant Alternatives) property holds in the estimated multinomial logit model. The test checks consistency of the estimated parameters between the original model and the model with fewer choices which are part of the full choices. We observe that the null hypothesis that the IIA property holds is rejected for subsets of choices excluding Hotel C or Hotel D. Hence, in the following examples, we only consider the nested logit model.

3.3 Discussion on Validity of Data Set

In this subsection, we discuss validity of the data set used in this numerical example and also address its limitations. We have selected full-service hotels in front of Kyoto station for the choices. These four hotels are the only full-service hotels in this location equipped with multiple restaurants, lounges, banquet rooms, conference halls for MICE (Meetings, Incentives, Conventions, Exhibitions) and a wedding center. We note that Kyoto city is designated as a global MICE strategic city by Japan Tourism Agency and takes its tourism strategy aiming for the ripple effects by the MICE tourists. On the other hand, the other hotels in this location are either select-service hotels or limited-service hotels, whose target customers are business travelers or leisure tourists putting a high value on costs while expecting minimal services from the hotels. Here, we mean 'in front of Kyoto Station' by within 500 meter distance of Kyoto station. In fact, the four hotels are all located within two-minute walking distance of Kyoto station, which offers great convenience to tourists coming from other parts of Japan by the superexpress train or other countries via Tokyo. Hotel C is in the terminal building, Hotel A is adjacent to the east side of the station, and Hotel B and Hotel D stand at the north or south front of the station.

In other words, we suppose that the target customers in this numerical analysis are the ones who look for a full-service hotel in front of Kyoto station, that is, aim to choose from the four hotels above. Among the four hotels, the price seems the primary characteristic that differentiates the hotel from the others and affects choice behaviors of customers. Hence, we set the price as the explanatory variable along with the holiday factor.

We remark that given the same location as in front of Kyoto Station, other attributes such as rating, customer reviews and capacity are inherent characteristics to each hotel, which remain unchanged for the short period in this analysis. Their effect on the random utility is reflected in the intercept α_i of $V_i^{(m)}$ in (4). Even if these characteristics are added as new explanatory variables for $V_i^{(m)}$, their coefficients are not reasonably estimated due to multicolinearity of the variables.

One way to observe these effects within our model is computing the choice probabilities when Hotel A, Hotel B, Hotel C and Hotel D offer the same prices. Although there is some impact from the price term, the choice probabilities, when the four room charges are the same, are considered to reflect the effect of the other characteristics on customers choice behaviors. Table 7 and 8 show the result, where we observe that for both a weekday and a day before a holiday, in all the cases, the ranking of choice probabilities is $D > C > A > B$. We note that if we focus on Hotel A, Hotel B and Hotel C, the ranking of the choice probabilities is the same as the order of ratings and customer reviews, which are examined in detail below. As for the choice probability of Hotel D, it is the highest in the rankings for both the weekday and the day before a holiday. Considering that Hotel D has the largest capacity as observed in Table 9 below and offers large number of accommodation plans as shown in Table 2, the capacity, promotion and advertisement are also considered to be the characteristics that affect the choice behavior of the customers.

Room charge	Hotel A	Hotel B	Hotel C	Hotel D
10,000	13%	5%	32%	50%
15,000	12%	4%	34%	49%
20,000	11%	4%	37%	48%
25,000	11%	3%	40%	46%
30,000	10%	3%	42%	45%

Table 7: Choice Probabilities of the Hotels for the same room charges, a weekday in November 2012.

Room charge	Hotel A	Hotel B	Hotel C	Hotel D
10,000	27%	2%	28%	43%
15,000	25%	2%	31%	43%
20,000	23%	2%	33%	42%
25,000	22%	2%	36%	41%
30,000	20%	1%	38%	40%

Table 8: Choice Probabilities of the Hotels for the same room charges, a day before a holiday in November 2012.

Next, Table 9 shows the summary of ratings, customer reviews and capacity for the four hotels. Although there is no official rating agency for Japanese hotels in Japan, Michelin Guide selects some excellent hotels in Kyoto and gives them its own rating. We can also refer to the ratings by online booking websites such as Booking.com and Expedia, though only the latest ratings are available. Then, we observe that Hotel C and Hotel D earn 4 pavilions and 2 pavilions respectively in Michelin Guide 2012. The ranking of the ratings is $C > A = D > B$ in Booking.com and $C > A > D = B$ in Expedia. We refer to Rakuten Travel for customer reviews, since the customers in the data set are the ones who booked through a Japanese booking website. We have aggregated the customer reviews in Rakuten Travel made for stays in 2012 for the four hotels and taken average of the ratings in the reviews. Then, the ranking of the customer reviews is $C > A > D > B$. As for capacity, the ranking of the total number of guest rooms is $D > C > A > B$. These numbers are as of April 2016 and unchanged from 2012.

	Hotel A	Hotel B	Hotel C	Hotel D
rating, Michelin Guide 2012	N/A	N/A	4	2
rating, Booking.com as of April 2016	4	3	5	4
rating, Expedia as of April 2016	4.0	3.5	4.5	3.5
customer reviews, Rakuten Travel as of 2012	4.27	4.17	4.65	4.25
total number of guest rooms	220	160	535	988

Table 9: Rating, Customer Reviews and Capacity for Hotel A, Hotel B, Hotel C and Hotel D.

Finally, we discuss the limitations of the data set: First, the customer type may be possibly different from what we have assumed. For instance, some customers who only care about the location may choose a hotel from a wider range of hotels in front of the station regardless of their service type. Second, although we have used the data as of November in 2012, this short period data do not reflect the change in the attributes other than the price. In order to analyze the effect of these attributes on choice behaviors of customers more thoroughly, the longer period data would be more appropriate. Third, the accuracy of the parameters estimation depends on the data quality. We have made the assumptions as in Section 3.1 in order to estimate the number of rooms booked and the quantity of inventory from the original online booking data. If we have full information on these items which each individual hotel has, the accuracy of the parameters estimation improves. However, it is also the advantage of the data set that hotels can maximize their expected sales by utilizing this data on the rival hotels observable at booking websites.

3.4 Estimated Optimal Room Charge and Expected Sales

In this subsection, we show the optimal room charges with the nested logit model in the cases of a normal weekday and a day before a holiday for the four hotels. Particularly, this result is obtained by the previously explained method with assumptions for construction of the data set described in Section 3.1. We also provide examples of the equilibrium room charges where every hotel maximizes its expected sales simultaneously. First, we show an example of the optimal room charge for a weekday. Hereafter we abbreviate the suffix m on the parameters for the notational simplicity. In addition to the estimated parameters of the nested logit model in Section 3.2, we assume the following: $\lambda = 6.07,\ x_1 = 11{,}985, x_2 = 11{,}000, x_3 = 19{,}938, x_4 = 18{,}000, T = 14, q_1 = \infty, q_2 = 13, q_3 = \infty, q_4 = \infty$. They are estimated from the data on the check-in date of a weekday in November 2012, when 85 rooms were booked over the 14 booking dates, the representative room charges for the four hotels were 11,985, 11,000, 19,938, 18,000 for Hotel A, Hotel B, Hotel C, and Hotel D respectively, and 13 rooms were booked for Hotel B before it got fully occupied while the other hotels had some rooms available at the check-in dates. We set q_i as the number of rooms sold during the booking period for the hotels which got fully-occupied, otherwise we set $q_i = \infty$. Note that in reality, each hotel knows its own quantity of room-inventory q_i, but needs to predict the quantity of the other hotels. We estimated λ by the maximum likelihood method for the Poisson distribution, which means that we set λ as the total number of rooms booked for the four hotels divided by 14, the number of booking days for the check-in date. Note that in order to enable the scenario by scenario computation of the expected sales in (17), which includes summation of large number of scenarios with low probabilities, we initially rescale the number of rooms of the hotels by considering 20 rooms as one batch. Hence we use the parameter $\tilde{\lambda} = 0.304, \tilde{q}_1 = \infty, \tilde{q}_2 = 1, \tilde{q}_3 = \infty, \tilde{q}_4 = \infty$ where $\tilde{\lambda} = \frac{\lambda}{20}, \tilde{q}_i = [\frac{q_i}{20}] + 1,\ (i = 1, 2, 3, 4)$ in (17), where $[\cdot]$ stands for the Gauss symbol. After the computation, the expected sales is obtained by multiplying the result by 20.

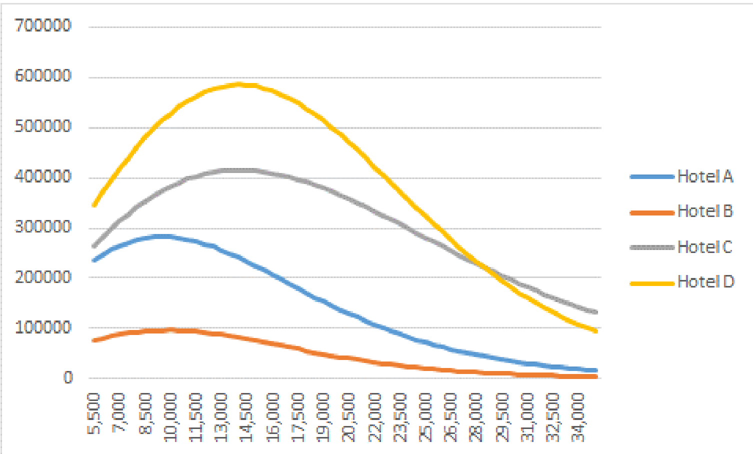

Figure 1: Relation between Estimated Expected Revenue and Room Charge (Individual Maximization Case), a weekday in November 2012, Hotel A, Hotel B, Hotel C and Hotel D.

Figure 1 shows the graphs of the expected sales as a function of the room charge for Hotel A, Hotel B, Hotel C, and Hotel D for the check-in date of a weekday in November 2012. We observe that as the room charge increases, the expected sales initially increases, and after the maximum point it turns to decreasing. The reason is that the higher room charge outweighs the decline in the expected number of rooms booked at first, and vice versa after the maximum point. These graphs also indicate the levels of the optimal room charges that maximize the expected sales of the hotels. (1) and (2) in Table 10 and 11 summarize the listed and optimal room charges, and their expected sales. We observe that Hotel A, Hotel C and Hotel D provided higher room charge than the estimated optimal level, while Hotel B offered a room charge close to the optimal level for this check-in date. We also find that each hotel potentially improves its estimated expected sales by setting the room charge at the optimal level.

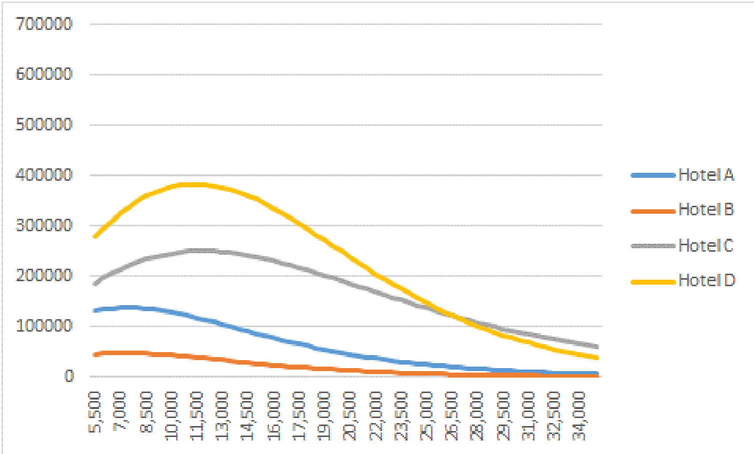

Figure 2: Relation between Estimated Expected Revenue and Room Charge (Equilibrium Case), a weekday in November 2012, Hotel A, Hotel B, Hotel C and Hotel D.

Figure 2 illustrates the expected sales of the hotels, when the other hotels set their room charges at the equilibrium prices. Here the equilibrium prices are the room charges of the hotels such that for each hotel, the price is at its optimal level given the other three hotels room charges at the equilibrium levels. In other words, they are the prices with which all the hotels maximize their expected sales simultaneously. The equilibrium room charge and the corresponding expected sales are (7,500, 137,873), (7,000, 47,112), (12,000, 250,380), (11,500, 383,706) for Hotel A, Hotel B, Hotel C, and Hotel D respectively. The equilibrium prices are obtained by iteration of the expected sales maximization where the room charge is replaced by the optimal one at each step. In detail, given the listed prices of the four hotels, we first obtain the optimal room charge of Hotel A. Then we replace the room charge of Hotel A by this, and calculate the optimal room charge of Hotel B. We repeat this maximization in the order of Hotel A, Hotel B, Hotel C, and Hotel D until the set of the four prices converges. Note that we have obtained the same equilibrium levels even if we switch the starting point of the iteration to Hotel B, Hotel C or Hotel D in this example as well as in the following example in Table 12 and 13, which shows the robustness of the obtained equilibrium prices at least around the individual optimization levels. Figure 1 and 2 show that in the equilibrium case, the expected sales are substantially reduced from the individual maximization. This implies that when all the hotels optimize their room charge simultaneously, as a result of the price competition, the expected sales fairly declined.

Table 10 and 11 summarize the optimal room charges and the expected sales for the individual maximization and the equilibrium cases. Compared to the individual maximization case, the equilibrium room charges are all at lower levels for this check-in date. It is also observed that the expected sales in the equilibrium case are lower than the individual maximization case by 34% to 51%, and than the listed case by 29% to 51%.

	Hotel A	Hotel B	Hotel C	Hotel D
(1) Listed	11,985	11,000	19,938	18,000
(2) Individual Maximization	9,500	10,000	14,000	14,000
(3) Equilibrium	7,500	7,000	12,000	11,500

Table 10: Room Charges for Listed, Individual Maximization, and Equilibrium, a weekday in November 2012.

	Hotel A	Hotel B	Hotel C	Hotel D
(1) Listed	268,749	95,454	366,263	537,706
(2) Individual Maximization	283,165	96,743	415,515	584,968
(3) Equilibrium	137,873	47,112	250,380	383,706

Table 11: Estimated Expected Sales for Listed, Individual Maximization, and Equilibrium, a weekday in November 2012.

The second example provides the case of a day before a holiday for the check-in date, when tourists stay for sightseeing for the weekend. Here we assume the following parameters: $\lambda = 3.93,\ x_1 = 18{,}036, x_2 = 17{,}771, x_3 = 26{,}400, x_4 = 20{,}000, T = 14, q_1 = \infty, q_2 = 7, q_3 = 29, q_4 = \infty$. Similarly to the previous example, they are estimated from the booking data on a day before a holiday in November 2012, when 55 rooms were booked in total for the four hotels in 14 booking days, Hotel B and Hotel C became fully occupied after 7 and 29 rooms were booked respectively, while the other hotels still had some rooms available at the check-in date, and 18,036, 177,771, 26,400, 20,000 were the representative room charges for Hotel A, Hotel B, Hotel C and Hotel D. As in the first example, we rescale the room number by regarding 20 rooms as one batch and multiply the number to the computational result. Hence we use $\tilde{\lambda} = 0.196, \tilde{q}_1 = \infty, \tilde{q}_2 = 1, \tilde{q}_3 = 2, \tilde{q}_4 = \infty$ for the computation of (17).

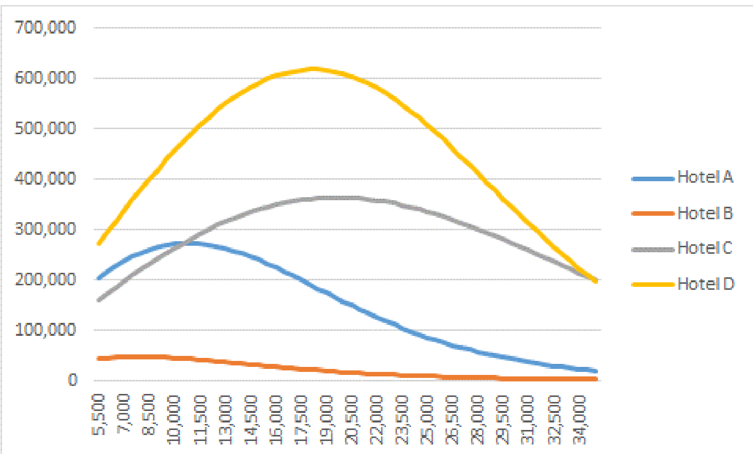

Figure 3: Relation between Estimated Expected Revenue and Room Charge (Individual Maximization Case), a day before a holiday in November 2012, Hotel A, Hotel B, Hotel C and Hotel D.

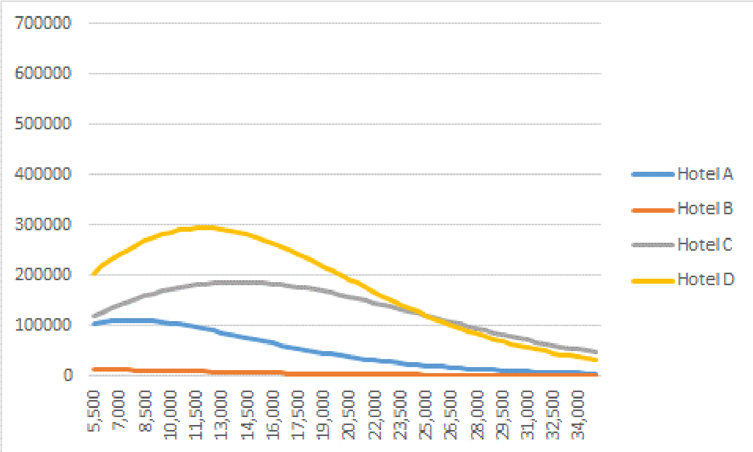

Figure 4: Relation between Estimated Expected Revenue and Room Charge (Equilibrium Case), a day before a holiday in November 2012, Hotel A, Hotel B, Hotel C and Hotel D.

Figure 3 and 4 illustrate the relation between the expected sales and the room charge for Hotel A, Hotel B, Hotel C, and Hotel D for both the individual maximization and the equilibrium cases. Table 12 displays the optimal room charges in the cases of the individual maximization and the equilibrium, as well as the listed room charges, and Table 13 shows their corresponding expected sales.

	Hotel A	Hotel B	Hotel C	Hotel D
Listed	18,036	17,771	26,400	20,000
Individual Maximization	11,000	8,000	19,500	18,000
Equilibrium	7,500	6,000	14,000	12,000

Table 12: Room Charges for Listed, Individual Maximization, and Equilibrium, a day before a holiday in November 2012.

	Hotel A	Hotel B	Hotel C	Hotel D
Listed	190,464	22,277	320,845	610,295
Individual Maximization	272,844	46,682	363,768	618,496
Equilibrium	110,506	11,662	186,397	293,878

Table 13: Estimated Expected Sales for Listed, Individual Maximization, and Equilibrium, a day before a holiday in November 2012.

The numerical results in the second example, show that the room charges in the equilibrium case are all lower than those in the individual maximization. Table 12 indicates that in the equilibrium situation where all the hotels aim to maximize their sales simultaneously, the room charges settle at low levels as a result of the price competition, which also leads to the substantially lower expected sales as in Table 13. It is observed that the expected sales in the equilibrium case are lower than those in the individual maximization case by 49% to 75%, and than in the listed case by 42% to 52%.

We can observe for both examples in Table 11 and 13 that expected sales with the listed room charge are in between those with the individual maximization and equilibrium prices. This may suggest that the hotels avoid lower expected sales caused by price competition which arises from the individual maximization of the hotels. On the other hand, the results also indicate that the hotels can increase their expected sales by setting their room charges at their estimated optimal levels by the individual maximization as long as the other hotels do not change their room charges to their optimal levels.

4 Conclusions and Future Research

This paper has analyzed online booking data of Kyoto, a city of international tourism that has 17 World Heritage Sites, for the first time in the literatures of hotel revenue management. The revenue management model used in this study reflects unique features of Japanese booking websites, fully transparent booking systems and limitation of the numbers of room available for booking. Firstly, this study has applied a quantitative revenue management model for estimates of choice probabilities of hotels by customers in online booking systems, which depend on room charges and types of a check-in date of hotels. The parameters in the model are estimated from a data set based on actual online booking data of customers for major four hotels in Kyoto city with application of econometric models, such as the multinomial and nested logit models. We have inferred the actual numbers of rooms booked under some appropriate assumptions, since these are not available in the original data set. This inference is meaningful because by extracting the information on rival hotels from data openly available on the web and using the model, hotels are able to predict their expected sales and optimal room charge. We have predicted optimal room charges and expected sales of the hotels when the other hotels' room charges are fixed or the other hotels also simultaneously maximize their expected sales, which is clearly useful for hotel managers. Moreover, this model enables hotel investors, such as hotel REIT and hotel funds to evaluate business value of hotels.

Finally, examining choice behaviors of customers in a longer period will be our future research topic: Particularly, we consider the other characteristics that do not change in a short term as explanatory variables, and we also take a wider range of hotels as the candidates for the choices.

Although we have used the data as of 2012, which was the year before the number of overseas tourists started to rise, investigating how the choice behaviors of customers have changed by the increase of the overseas tourists resulting from JPY weakening, is also our next research topic.

5 Acknowledgments

We are grateful to the editor-in-chief, Professor Abraham Pizam, the managing editor, Professor Manuel Rivera and two anonymous referees for their precious comments, which substantially improve the previous version of this paper. We also thank Professor Naoyuki Yoshino, Professor Nobuo Takahashi, Professor Takashi Obinata, Professor

Masaki Yoneyama and Professor Seisho Sato for their valuable suggestions. We are grateful to National Institute of Informatics for offering data used in this study and to Kotaro Urakami for organizing the database. This research is supported by JSPS KAKENHI Grant Numbers 25380389 and 26283019.

References

[1] G.Abrate, G.Fraquelli, G.Viglia. *Dynamic pricing strategies: Evidence from European hotels*, Int. J. Hospitality Management, 31.1 (2012) 160-168.

[2] T.Amemiya: *Advanced econometrics*, Harvard university press, (1985).

[3] C.Anderson, X.Xie: *A Choice-Based Dynamic Programming Approach for Setting Opaque Prices*, Production and Operations Management 21 (2012) 590-605.

[4] R.Badinelli: *An Optimal, Dynamic Policy for Hotel Yield Management*, European Journal of Operations Research 121 (2000) 476-503.

[5] T.Baker, D.Collier: *A Comparative Revenue Analysis of Hotel Yield Management Heuristics*, Decision Sciences, 30 (1999) 239-263.

[6] G.Bitran, S.Mondschein: *An Application of Yield Management to the Hotel Industry Considering Multiple Day Stays*, Operations Research 43 (1995) 427-443.

[7] L.V.Casaló, C.Flavián, M.Guinalíu, Y.Ekinci: *Do online hotel rating schemes influence booking behaviors ?*, Int. J. Hospitality Management, 49 (2015) 28-36.

[8] K.Donaghy, U.McMahon, D.McDowell: *Yield Management: an Overview*, Int. J. Hospitality Management, 14 (1995) 139-150.

[9] M.Gibson: *Understanding the risk of synthetic CDOs*, Finance and Economics Discussion Series, 36 (2004).

[10] R.Hanks, R.Cross, R.Noland: *Discounting in the Hotel Industry: A New Approach*, The Cornell H.R.A Quarterly (1992).

[11] J.Hausman, D.McFadden: *Specification Tests for the Multinomial Logit Model*, Econometrica, 52 (1984) 1219-1240.

[12] J.Hull, A.White: *The Risk of Tranches Created from Mortgages*, Financial Analysts Journal, 66 (2010) 54-67.

[13] S.Kimes: Yield Management: *A Tool for Capacity-Constrained Service Firms*, Journal of Operations Management, 8 (1989) 348-363.

[14] S.Kimes: *The Basics of Yield Management*: The Cornell H.R.A Quarterly (1989).

[15] R.Ladhari, M.Michaud: *eWOM Effects on Hotel Booking Intentions, Attitudes, Trust, and Website Perceptions*, Int. J. Hospitality Management, 46 (2015) 36-45.

[16] J.N.K.Liu, E.Y.Zhang: *An Investigation of Factors Affecting Customer Selection of Online Hotel Booking Channels*, Int. J. Hospitality Management, 39 (2014) 71-83.

[17] L.Masiero, C.Y.Heo, B.Pan. *Determining guests' willingness to pay for hotel room attributes with a discrete choice model*, Int. J. Hospitality Management, 49 (2015) 117-124.

[18] L.Masiero, and J.L.Nicolau. *Price sensitivity to tourism activities: looking for determinant factors*, Tourism Economics, 18.4 (2012) 675-689.

[19] L.Masiero, J.L.Nicolau, R.Law. *A demand-driven analysis of tourist accommodation price: A quantile regression of room bookings*, Int. J. Hospitality Management, 50 (2015) 1-8.

[20] L.Masiero, B.Pan, C.Y.Heo. *Asymmetric preference in hotel room choice and implications on revenue management*, Int. J. Hospitality Management, 56 (2016) 18-27.

[21] D.McFadden: *Modelling the Choice of Residential Location, Spatial Interaction Theory and Planning Models*, California: Institute of Transportation Studies, University of California, (1978) 75-96.

[22] D.McFadden: *Econometric Models of Probabilistic Choice*, Structural Analysis of Discrete Data and Econometric Applications, Cambridge, MA. MIT Press. (1981) 198-272.

[23] W.Lieberman: *Debunking the Myths of Yield Management*, The Cornell H.R.A Quarterly (1993).

[24] J.L.Nicolau, L.Masiero. *Relationship between price sensitivity and expenditures in the choice of tourism activities at the destination*, Tourism Economics, 19.1 (2013) 101-114.

[25] B.M.Noone, A.S.Mattila: *Hotel Revenue Management and the Internet: The Effect of Price Presentation Strategies on Customer's Willingness to Book*, Int. J. Hospitality Management, 28 (2009) 272-279.

[26] L.Weatherford, S.Bodily: *A Taxonomy and Research Overview of Perishable-Asset Revenue Management: Yield Management, Overbooking, and Pricing*, Operations Research, 40 (1992) 831-844.

[27] L.Wang, R.Law, B.D.Guillet, K.Hung, D.K.C.Fong: *Impact of Hotel Website Quality on Online Booking Intentions: eTrust as a Mediator*, Int. J. Hospitality Management, 47 (2015) 108-115.

[28] G.Viglia, R.Furlan, A.Ladron-de-Guevara. *Please, talk about it! When hotel popularity boosts preferences*, Int. J. Hospitality Management, 42 (2014) 155-164.

[29] G.Viglia, A.Mauri, M.Carricano. *The exploration of hotel reference prices under dynamic pricing scenarios and different forms of competition*, Int. J. Hospitality Management, 52 (2016) 46-55.

A Some properties of the equation (8) with respect to ν_k

Proposition 1. *Let $\tilde{C}_k = \{j \in C_k | \gamma_j = 1\}$. Suppose that $\tilde{C}_k$ is nonempty. Then we have*

$$\lim_{\nu_k \to 0} \left\{ \sum_{j \in C_k} \exp\left(\frac{V_j^{(m)}}{\nu_k}\right) 1_{\{\gamma_j = 1\}} \right\}^{\nu_k} = \max_{j \in \tilde{C}_k} \exp(V_j^{(m)}). \tag{26}$$

Moreover,

$$\left\{ \sum_{j \in C_k} \exp\left(\frac{V_j^{(m)}}{\nu_k}\right) 1_{\{\gamma_j = 1\}} \right\}^{\nu_k} \tag{27}$$

is nondecreasing on $0 < \nu_k \leq 1$.

Proof. First note that

$$\max_{j \in \tilde{C}_k} \exp\left(\frac{V_j^{(m)}}{\nu_k}\right) \leq \sum_{j \in \tilde{C}_k} \exp\left(\frac{V_j^{(m)}}{\nu_k}\right) \leq \#\tilde{C}_k \max_{j \in \tilde{C}_k} \exp\left(\frac{V_j^{(m)}}{\nu_k}\right). \tag{28}$$

Then we have

$$\max_{j \in \tilde{C}_k} \exp(V_j^{(m)}) \leq \left\{ \sum_{j \in \tilde{C}_k} \exp\left(\frac{V_j^{(m)}}{\nu_k}\right) \right\}^{\nu_k} \leq \left(\#\tilde{C}_k\right)^{\nu_k} \max_{j \in \tilde{C}_k} \exp(V_j^{(m)}). \tag{29}$$

Taking the limit as $\nu_k \to 0$, we obtain the desired result.

Next, we let $0 < \nu_k < \tilde{\nu}_k \leq 1$. Then we have

$$\begin{aligned}
\frac{\left\{ \sum_{j \in \tilde{C}_k} \exp\left(\frac{V_j^{(m)}}{\nu_k}\right) \right\}^{\nu_k}}{\left\{ \sum_{j \in \tilde{C}_k} \exp\left(\frac{V_j^{(m)}}{\tilde{\nu}_k}\right) \right\}^{\tilde{\nu}_k}} &= \left(\sum_{j \in \tilde{C}_k} \frac{\exp\left(\frac{V_j^{(m)}}{\nu_k}\right)}{\left\{ \sum_{j \in \tilde{C}_k} \exp\left(\frac{V_j^{(m)}}{\tilde{\nu}_k}\right) \right\}^{\frac{\tilde{\nu}_k}{\nu_k}}} \right)^{\nu_k} \\
&= \left(\sum_{j \in \tilde{C}_k} \left\{ \frac{\exp\left(\frac{V_j^{(m)}}{\tilde{\nu}_k}\right)}{\sum_{j \in \tilde{C}_k} \exp\left(\frac{V_j^{(m)}}{\tilde{\nu}_k}\right)} \right\}^{\frac{\tilde{\nu}_k}{\nu_k}} \right)^{\nu_k} \\
&\leq \left(\sum_{j \in \tilde{C}_k} \frac{\exp\left(\frac{V_j^{(m)}}{\tilde{\nu}_k}\right)}{\sum_{j \in \tilde{C}_k} \exp\left(\frac{V_j^{(m)}}{\tilde{\nu}_k}\right)} \right)^{\nu_k} \\
&= 1.
\end{aligned} \tag{30}$$

Hence

$$\left\{ \sum_{j \in C_k} \exp\left(\frac{V_j^{(m)}}{\nu_k}\right) 1_{\{\gamma_j = 1\}} \right\}^{\nu_k} \tag{31}$$

is nondecreasing on $0 < \nu_k \leq 1$. □

B Proof of Theorem 1

In this appendix, we give the proof of Theorem 1. First, we show the following lemma on the upper bound estimation on the conditional expectation on the number of rooms booked.

Lemma 2.

$$\mathbf{E}[R_T^i|N_T^{(m)}=k] \le k\left(\max_{l\in C_{k_i},l\neq i}\left\{\frac{1}{1+\exp(\frac{V_l^{(m)}-V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})}\right\}+\max_{l\notin C_{k_i}}\left\{\frac{1}{1+\exp(V_l^{(m)}-V_i^{(m)}(x_i^{(m)}))}\right\}\right). \tag{32}$$

Proof. Noting that there exists $j\neq i$ such that $q_j^{(m)}=\infty$, by (15), (16), we have

$$\begin{aligned}
&\mathbf{E}[R_T^i|N_T^{(m)}=k]\\
&=\sum_{(i_1,\dots,i_k)\in S_k} a^i_{(i_1,\dots,i_k)} p_{i_1}^{(m)\gamma^1_{(i_1,\dots,i_k)}}\dots p_{i_k}^{(m)\gamma^k_{(i_1,\dots,i_k)}}\\
&=\sum_{l=1}^k\sum_{(i_1,\dots,i_k)\in S_k} 1_{\{i_l=i\}} p_{i_1}^{(m)\gamma^1_{(i_1,\dots,i_k)}}\dots p_{i_k}^{(m)\gamma^k_{(i_1,\dots,i_k)}}\\
&=\sum_{l=1}^k\sum_{(i_1,\dots,i_{l-1},i,i_{l+1},\dots,i_k)\in S_k} p_{i_1}^{(m)\gamma^1_{(i_1,\dots,i_{l-1},i,i_{l+1},\dots,i_k)}}\dots p_{i}^{(m)\gamma^l_{(i_1,\dots,i_{l-1},i,i_{l+1},\dots,i_k)}}\dots p_{i_k}^{(m)\gamma^k_{(i_1,\dots,i_{l-1},i,i_{l+1},\dots,i_k)}}\\
&\le\sum_{l=1}^k\max_{\gamma\in\Gamma,\gamma_j=1} p_i^{(m)\gamma}\sum_{(i_1,\dots,i_{l-1},i,i_{l+1},\dots,i_k)\in S_k}\\
&\quad p_{i_1}^{(m)\gamma^1_{(i_1,\dots,i_{l-1},i,i_{l+1},\dots,i_k)}}\dots p_{i_{l-1}}^{(m)\gamma^{l-1}_{(i_1,\dots,i_{l-1},i,i_{l+1},\dots,i_k)}} p_{i_{l+1}}^{(m)\gamma^{l+1}_{(i_1,\dots,i_{l-1},i,i_{l+1},\dots,i_k)}}\dots p_{i_k}^{(m)\gamma^k_{(i_1,\dots,i_{l-1},i,i_{l+1},\dots,i_k)}}\\
&=k\max_{\gamma\in\Gamma,\gamma_j=1} p_i^{(m)\gamma}.
\end{aligned}\tag{33}$$

Here we have used the equality

$$
\begin{aligned}
&\sum_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)\in S_k} \\
&p_{i_1}^{(m)\gamma^1_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)}} \ldots p_{i_{l-1}}^{(m)\gamma^{l-1}_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)}} p_{i_{l+1}}^{(m)\gamma^{l+1}_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)}} \ldots p_{i_k}^{(m)\gamma^k_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)}} \\
&= \sum_{i_1=1}^{L} \cdots \sum_{i_{l-1}=1}^{L} \sum_{i_{l+1}=1}^{L} \cdots \sum_{i_k=1}^{L} \\
&p_{i_1}^{(m)\gamma^1_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)}} \ldots p_{i_{l-1}}^{(m)\gamma^{l-1}_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)}} p_{i_{l+1}}^{(m)\gamma^{l+1}_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)}} \ldots p_{i_k}^{(m)\gamma^k_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)}} \\
&= \sum_{i_1=1}^{L} \cdots \sum_{i_{l-1}=1}^{L} \sum_{i_{l+1}=1}^{L} \cdots \sum_{i_{k-1}=1}^{L} \\
&p_{i_1}^{(m)\gamma^1_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)}} \ldots p_{i_{l-1}}^{(m)\gamma^{l-1}_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)}} p_{i_{l+1}}^{(m)\gamma^{l+1}_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)}} \ldots p_{i_{k-1}}^{(m)\gamma^{k-1}_{(i_1,\ldots,i_{l-1},i,i_{l+1},\ldots,i_k)}} \\
&\ldots \\
&= 1, \qquad (34)
\end{aligned}
$$

which follows from (3).

(i) The case $j \in C_{k_i}$. In this case, since $p_i^{(m)\gamma}$ in (8) is maximized at $\gamma \in \Gamma$ such that $\gamma_i = \gamma_j = 1, \gamma_l = 0$ $(l \neq i, j)$,

$$
\begin{aligned}
\max_{\gamma\in\Gamma,\gamma_j=1} p_i^{\gamma} &= 1 \cdot \frac{\exp(\frac{V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})}{\exp(\frac{V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}}) + \exp(\frac{V_j^{(m)}}{\nu_{k_i}})} \\
&= \frac{1}{1+\exp(\frac{V_j^{(m)}-V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})}. \qquad (35)
\end{aligned}
$$

(ii) The case $j \notin C_{k_i}$. Let k_j be the index in $\{1,\ldots,n\}$ such that $j \in C_{k_j}$. In this case, for all $\gamma \in \Gamma, \gamma_j = 1$,

$$
p_i^{(m)\gamma} \leq \frac{\{\sum_{l\in C_{k_i}} \exp(\frac{V_l^{(m)}}{\nu_{k_i}})1_{\{\gamma_l=1\}}\}^{\nu_{k_i}}}{\{\sum_{l\in C_{k_i}} \exp(\frac{V_l^{(m)}}{\nu_{k_i}})1_{\{\gamma_l=1\}}\}^{\nu_{k_i}} + \exp(\frac{V_j^{(m)}}{\nu_{k_j}})^{\nu_{k_j}}} \cdot \frac{\exp(\frac{V_i^{(m)}}{\nu_{k_i}})1_{\{\gamma_i=1\}}}{\sum_{l\in C_{k_i}} \exp(\frac{V_l^{(m)}}{\nu_{k_i}})1_{\{\gamma_l=1\}}}. \qquad (36)
$$

When $\gamma_l = 0$ for all $l \in C_{k_i}, l \neq i$,

$$
\begin{aligned}
&\frac{\{\sum_{l\in C_{k_i}} \exp(\frac{V_l^{(m)}}{\nu_{k_i}})1_{\{\gamma_l=1\}}\}^{\nu_{k_i}}}{\{\sum_{l\in C_{k_i}} \exp(\frac{V_l^{(m)}}{\nu_{k_i}})1_{\{\gamma_l=1\}}\}^{\nu_{k_i}} + \exp(\frac{V_j^{(m)}}{\nu_{k_j}})^{\nu_{k_j}}} \cdot \frac{\exp(\frac{V_i^{(m)}}{\nu_{k_i}})1_{\{\gamma_i=1\}}}{\sum_{l\in C_{k_i}} \exp(\frac{V_l^{(m)}}{\nu_{k_i}})1_{\{\gamma_l=1\}}} \\
&\leq \frac{\exp(\frac{V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})^{\nu_{k_i}}}{\exp(\frac{V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})^{\nu_{k_i}} + \exp(\frac{V_j^{(m)}}{\nu_{k_j}})^{\nu_{k_j}}} \cdot 1 \\
&= \frac{1}{1+\exp(V_j^{(m)} - V_i^{(m)}(x_i^{(m)}))}. \qquad (37)
\end{aligned}
$$

When there exists $l' \in C_{k_i}, l' \neq i$ such that $\gamma_{l'} = 1$,

$$\begin{aligned}
&\frac{\{\sum_{l\in C_{k_i}} \exp(\frac{V_l^{(m)}}{\nu_{k_i}})1_{\{\gamma_l=1\}}\}^{\nu_{k_i}}}{\{\sum_{l\in C_{k_i}} \exp(\frac{V_l^{(m)}}{\nu_{k_i}})1_{\{\gamma_l=1\}}\}^{\nu_{k_i}} + \exp(\frac{V_j^{(m)}}{\nu_{k_j}})^{\nu_{k_j}}} \cdot \frac{\exp(\frac{V_i^{(m)}}{\nu_{k_i}})1_{\{\gamma_i=1\}}}{\sum_{l\in C_{k_i}} \exp(\frac{V_l^{(m)}}{\nu_{k_i}})1_{\{\gamma_l=1\}}} \\
&\leq 1 \cdot \frac{\exp(\frac{V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})}{\exp(\frac{V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}}) + \exp(\frac{V_{l'}^{(m)}}{\nu_{k_i}})} \\
&\leq \max_{l\in C_{k_i}, l\neq i} \left(\frac{1}{1+\exp(\frac{V_l^{(m)} - V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})} \right).
\end{aligned} \tag{38}$$

Hence

$$\max_{\gamma\in\Gamma,\gamma_j=1} p_i^{(m)\gamma} \leq \max\left(\max_{l\in C_{k_i}, l\neq i} \left(\frac{1}{1+\exp(\frac{V_l^{(m)} - V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})} \right), \frac{1}{1+\exp(V_j^{(m)} - V_i^{(m)}(x_i^{(m)}))} \right). \tag{39}$$

Thus we have

$$\max_{\gamma\in\Gamma,\gamma_j=1} p_i^{\gamma} \leq \max_{l\in C_{k_i}, l\neq i} \left\{ \frac{1}{1+\exp(\frac{V_l^{(m)} - V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})} \right\} + \max_{l\notin C_{k_i}} \left\{ \frac{1}{1+\exp(V_l^{(m)} - V_i^{(m)}(x_i^{(m)}))} \right\} \tag{40}$$

and the proof is complete. $\square$

Then we give the proof of Theorem 1.

Let

$$f(x_i^{(m)}) = x_i^{(m)} \sum_{k=0}^{\infty} \mathbf{E}[R_T^i | N_T^{(m)} = k] \exp(-\lambda^{(m)}T) \frac{(\lambda^{(m)}T)^k}{k!}, \tag{41}$$

$$f_N(x_i^{(m)}) = x_i^{(m)} \sum_{k=0}^{N} \mathbf{E}[R_T^i | N_T^{(m)} = k] \exp(-\lambda^{(m)}T) \frac{(\lambda^{(m)}T)^k}{k!}. \tag{42}$$

Note that $f_N(x_i^{(m)})$ is a continuous function on $x_i^{(m)} \in [0,\infty)$ by (15).

First we show that $f_N(x_i^{(m)})$ converges to $f(x_i^{(m)})$ uniformly on $[0,\infty)$ as $N \to \infty$. By

Lemma 2,

$$
\begin{aligned}
&\sup_{x_i^{(m)}\in[0,\infty)} \left| f_N(x_i^{(m)}) - f(x_i^{(m)}) \right| \\
&= \sup_{x_i^{(m)}\in[0,\infty)} \left| x_i^{(m)} \sum_{k=N+1}^{\infty} \mathbf{E}[R_T^i | N_T^{(m)} = k] \exp(-\lambda^{(m)}T) \frac{(\lambda^{(m)}T)^k}{k!} \right| \\
&\leq \sup_{x_i^{(m)}\in[0,\infty)} \left| \sum_{k=N+1}^{\infty} k \left(\max_{l\in C_{k_i}, l\neq i} \left\{ \frac{x_i^{(m)}}{1+\exp(\frac{V_l^{(m)}-V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})} \right\} + \max_{l\notin C_{k_i}} \left\{ \frac{x_i^{(m)}}{1+\exp(V_l^{(m)}-V_i^{(m)}(x_i^{(m)}))} \right\} \right) \right. \\
&\quad \left. \times \exp(-\lambda^{(m)}T) \frac{(\lambda^{(m)}T)^k}{k!} \right| \\
&\leq M \left| \sum_{k=N+1}^{\infty} k \exp(-\lambda^{(m)}T) \frac{(\lambda^{(m)}T)^k}{k!} \right| \\
&\to 0 \ (N\to\infty),
\end{aligned} \tag{43}
$$

where

$$
M = \sup_{x_i^{(m)}\in[0,\infty)} \left(\max_{l\in C_{k_i}, l\neq i} \left\{ \frac{x_i^{(m)}}{1+\exp(\frac{V_l^{(m)}-V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})} \right\} + \max_{l\notin C_{k_i}} \left\{ \frac{x_i^{(m)}}{1+\exp(V_l^{(m)}-V_i^{(m)}(x_i^{(m)}))} \right\} \right) < \infty. \tag{44}
$$

Here we have used the fact that since $\beta_i < 0$,

$$
\lim_{x_i^{(m)}\to\infty} \left(\max_{l\in C_{k_i}, l\neq i} \left\{ \frac{x_i^{(m)}}{1+\exp(\frac{V_l^{(m)}-V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})} \right\} + \max_{l\notin C_{k_i}} \left\{ \frac{x_i^{(m)}}{1+\exp(V_l^{(m)}-V_i^{(m)}(x_i^{(m)}))} \right\} \right) = 0. \tag{45}
$$

Hence $f_N(x_i^{(m)})$ converges to $f(x_i^{(m)})$ uniformly on $[0,\infty)$ and $f(x_i^{(m)})$ is a continuous function on $x_i^{(m)} \in [0,\infty)$.

Similarly,

$$
\begin{aligned}
0 &\le f(x_i^{(m)}) \\
&= x_i^{(m)} \sum_{k=0}^{\infty} \mathbf{E}[R_T^i | N_T^{(m)} = k] \exp(-\lambda^{(m)} T) \frac{(\lambda^{(m)} T)^k}{k!} \\
&\le \left(\max_{l \in C_{k_i}, l \ne i} \left\{ \frac{x_i^{(m)}}{1 + \exp(\frac{V_l^{(m)} - V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})} \right\} + \max_{l \notin C_{k_i}} \left\{ \frac{x_i^{(m)}}{1 + \exp(V_l^{(m)} - V_i^{(m)}(x_i^{(m)}))} \right\} \right) \\
&\quad \times \sum_{k=0}^{\infty} k \exp(-\lambda^{(m)} T) \frac{(\lambda^{(m)} T)^k}{k!} \\
&= \lambda^{(m)} T \left(\max_{l \in C_{k_i}, l \ne i} \left\{ \frac{x_i^{(m)}}{1 + \exp(\frac{V_l^{(m)} - V_i^{(m)}(x_i^{(m)})}{\nu_{k_i}})} \right\} + \max_{l \notin C_{k_i}} \left\{ \frac{x_i^{(m)}}{1 + \exp(V_l^{(m)} - V_i^{(m)}(x_i^{(m)}))} \right\} \right),
\end{aligned}
\tag{46}
$$

and it follows that $\lim_{x_i^{(m)} \to \infty} f(x_i^{(m)}) = 0$.

Since $f(x_i^{(m)})$ is continuous, $f(x_i^{(m)})$ has a maximum point on any finite interval of the form $[0, K]$, and for any sufficiently small $\epsilon > 0$, we can take $K > 0$ such that for any $x_i^{(m)} > K$, $0 \le f(x_i^{(m)}) < \epsilon$. Therefore $f(x_i^{(m)})$ has a maximum point in $x_i^{(m)} \in [0, \infty)$. $\square$

C Proof of Lemma 1

Since the multinomial logit model is the special case of the nested logit model with $\nu_1 = \cdots = \nu_n = 1$, we only prove (20).

First note that for $j \in \{1, \ldots, L\}$,

$$
\begin{aligned}
\exp\left(\frac{V_j^{(m)}}{\nu_k} \right) 1_{\{\gamma_j = 1\}} &= \exp\left(\frac{\tilde{V}_j^{(m)} + \alpha_1 + \delta_1 y^{(m)}}{\nu_k} \right) 1_{\{\gamma_j = 1\}} \\
&= \exp\left(\frac{\alpha_1 + \delta_1 y^{(m)}}{\nu_k} \right) \exp\left(\frac{\tilde{V}_j^{(m)}}{\nu_k} \right) 1_{\{\gamma_j = 1\}}.
\end{aligned}
\tag{47}
$$

Therefore we have

$$
\begin{aligned}
p_i^{(m)\gamma} &= \frac{\{\sum_{j \in C_{k_i}} \exp(\frac{V_j^{(m)}}{\nu_{k_i}}) 1_{\{\gamma_j = 1\}}\}^{\nu_{k_i}}}{\sum_{k=1}^n \{\sum_{j \in C_k} \exp(\frac{V_j^{(m)}}{\nu_k}) 1_{\{\gamma_j = 1\}}\}^{\nu_k}} \cdot \frac{\exp(\frac{V_i^{(m)}}{\nu_{k_i}}) 1_{\{\gamma_i = 1\}}}{\sum_{j \in C_{k_i}} \exp(\frac{V_j^{(m)}}{\nu_{k_i}}) 1_{\{\gamma_j = 1\}}} \\
&= \frac{\exp(\alpha_1 + \delta_1 y^{(m)}) \{\sum_{j \in C_{k_i}} \exp(\frac{\tilde{V}_j^{(m)}}{\nu_{k_i}}) 1_{\{\gamma_j = 1\}}\}^{\nu_{k_i}}}{\exp(\alpha_1 + \delta_1 y^{(m)}) \sum_{k=1}^n \{\sum_{j \in C_k} \exp(\frac{\tilde{V}_j^{(m)}}{\nu_k}) 1_{\{\gamma_j = 1\}}\}^{\nu_k}} \cdot \frac{\exp(\frac{\alpha_1 + \delta_1 y^{(m)}}{\nu_{k_i}}) \exp(\frac{\tilde{V}_i^{(m)}}{\nu_{k_i}}) 1_{\{\gamma_i = 1\}}}{\exp(\frac{\alpha_1 + \delta_1 y^{(m)}}{\nu_{k_i}}) \sum_{j \in C_{k_i}} \exp(\frac{\tilde{V}_j^{(m)}}{\nu_{k_i}}) 1_{\{\gamma_j = 1\}}} \\
&= \frac{\{\sum_{j \in C_{k_i}} \exp(\frac{\tilde{V}_j^{(m)}}{\nu_{k_i}}) 1_{\{\gamma_j = 1\}}\}^{\nu_{k_i}}}{\sum_{k=1}^n \{\sum_{j \in C_k} \exp(\frac{\tilde{V}_j^{(m)}}{\nu_k}) 1_{\{\gamma_j = 1\}}\}^{\nu_k}} \cdot \frac{\exp(\frac{\tilde{V}_i^{(m)}}{\nu_{k_i}}) 1_{\{\gamma_i = 1\}}}{\sum_{j \in C_{k_i}} \exp(\frac{\tilde{V}_j^{(m)}}{\nu_{k_i}}) 1_{\{\gamma_j = 1\}}}. \quad \square
\end{aligned}
\tag{48}
$$