Application of Online Booking Data to Hotel Revenue Management

Taiga Saito[a], Akihiko Takahashi[b], Noriaki Koide[c], Yu Ichifuji[d]

[a]*Graduate School of Economics, The University of Tokyo*
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[b]*Graduate School of Economics, The University of Tokyo*
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[c]*Joint Support-Center for Data Science Research, Research Organization of Information and systems*
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430
[d]*Center for Information and Communication Technology, Nagasaki University*
1-14, Bunkyo-machi, Nagasaki City, Nagasaki 852-8521

Abstract

This paper presents an application of online booking data, comprised of big data crawled from a hotel booking website to hotel revenue management. It is important to build a quantitative revenue management method for online hotel booking systems incorporating overbooking strategies, because of increasing numbers of bookings through online booking websites and last-minute cancellations, which cause serious damage to hotel management. We construct a quantitative overbooking model for online booking systems combined with customers' choice behaviors estimated from the data. Firstly, we present the overbooking model for online booking systems. Secondly, we estimate the choice behaviors of the customers from the online booking data by a discrete choice model. Thirdly, combining the estimated discrete choice model with the theoretical overbooking model, we investigate the expected sales maximization problem where we numerically solve the optimal overbooking level and room charge. Finally, we provide numerical examples of the optimal overbooking strategies and room charges using online booking data of two major luxury hotels in Shinjuku ward, Tokyo. This method, which utilizes online booking data available by crawling from booking websites, helps hotels obtain an optimal room charge and overbooking level maximizing the expected sales.

Keywords: Online hotel booking data, Choice behaviors, Revenue management, Overbooking strategies

1. Introduction

In this paper, we present an application of online hotel booking data, which are big data crawled from an online booking website, to revenue management of hotels by a quantitative overbooking model for online booking systems. Firstly, we estimate the customers' choice behaviors in the online booking system from the online booking data by a discrete choice model. Secondly, we build a quantitative revenue management model under overbooking that incorporates the last-minute cancellations and the overselling costs that hotels incur when the number of customers appearing on the check-in date exceeds the hotels' capacity. Then, combining the choice behaviors estimated from the big data and the theoretical overbooking model, we investigate an overbooking strategy that maximizes expected sales of a certain room type of a hotel in an online booking system, where customers choose a hotel from a group comparing the room charge with those of the rival hotels. Furthermore, we provide numerical examples using online booking data of two major luxury hotels in Shinjuku, Tokyo, which are crawled from a Japanese online booking website. In addition, we investigate a Nash equilibrium for the room charge and expected sales of the hotels. The revenue management model under overbooking helps hotels determine the overbooking strategy, which maximizes the expected sales in an online booking system.

Email addresses: staiga@e.u-tokyo.ac.jp (Taiga Saito), akihikot@e.u-tokyo.ac.jp (Akihiko Takahashi), n_koide@nii.ac.jp (Noriaki Koide), ichifuji@nagasaki-u.ac.jp (Yu Ichifuji)

Owing to the advancement in information technologies, an increasing number of hotel customers are booking rooms through online booking systems. They can conveniently reserve rooms by comparing prices and conditions of other hotels. Under these circumstances, the hotel revenue management for online booking systems, where hotels set the room charges at the optimal levels to maximize their expected revenues by considering the room charges of the rival hotels, is becoming increasingly important.

As the share of the rooms booked through online booking systems increases, the number of last-minute cancellations is also rising due to the convenience of the booking (Walls (2016), Delgado (2016), Horiguchi (2015)). For example, a pattern of cancellation called "juggling," in which customers reserve rooms for multiple hotels, then choose one room/hotel and cancel the others at the last minute, represents a critical problem for hotels that sell rooms through online booking systems (Walls (2016), Mandelbaum (2016), Takizawa (2017)). Such last-minute cancellations lead to revenue losses for hotels.

Although overbooking is practiced to cope with cancellations, it is done by rule of thumb due to the following difficulties. If the number of cancellations and no-shows is more than expected and these generate some vacant rooms, the hotel incurs some loss on sales. On the other hand, if the number of cancellations and no-shows is less than expected, the hotel gets into a situation of over-sale. In such a case, the hotel needs to displace the guests to other hotels by paying for the room charge and offering extra compensation if necessary, which are additional costs for the hotel (Toh (1986), Toh and Dekay (2002)). To deal with such an issue, we investigate the overbooking problem quantitatively together with the expected sales maximization for online booking systems.

For studies related to overbooking, Toh (1986) and Toh and Dekay (2002) study practices of no-shows, late cancellations, overbooking and over-sale. Dekay et al. (2004) examine penalties on hotel customers for cancellations and no-shows as well as practices of upgrading and walks in the case of over-sale. Chen et al. (2011) analyze the impact of cancellation fees on customer behaviors of searching for hotels by multinomial logit model. Chen and Xie (2013) investigate the cancellation policies of U.S. hotels.

For quantitative studies on overbooking models, Rothstein (1974) deals with an overbooking problem to determine an optimal booking policy using a Markovian sequential decision model. Liberman and Yechiali (1978) deal with the overbooking problem using stochastic cancellations. Gallego and Ryzin (1994) investigate formulations of revenue management including in cases of overbooking. Koide and Ishii (2005) consider a model for an optimal room allocation for early discount considering cancellations and overbooking. Sierag et al. (2015) consider an optimal choice of fare products to offer under overbooking. We summarize this study's contributions to the existing literature in Section 4.1.

As for applications of data on online hotel booking websites to hotel management and marketing, Ali et al. (2016) propose a classification technique for feature review's identification and opinion mining about hotels and hotel features. Kardaras et al. (2013) investigate content presentation and media adaptation of tourism websites. Korfiatis and Poulos (2013) examine the usage of online consumer reviews as a source for demographic recommendation. Shuai and Wu (2011) analyze the impact of internet marketing on hotel performance. Yu et al. (2017) design a mathematical model to select appropriate hotels on websites using online reviews of hotels with linguistic distribution assessments. Bilgihan et al. (2015) examine the concepts of online customer experiences and flow and their meditating roles in influencing customers' loyalty to a hotel booking website. Neirotti et al. (2016) analyze panel data on the value generation and appropriation mechanisms related to Italian hotels' online visibility on infomediary platforms. Xu et al. (2017) examine customer satisfaction and dissatisfaction toward attributes of hotel products and services based on online customer reviews. Hernandez-Ortega (2018) analyzes whether sociopsychological distance is the underlying mechanism that mediates the effect of online customer review aspects on the receiver's responses. Raguseo et al. (2017) explore how hotels can drive value their way when they manage their visibility on infomediary platforms. Huang et al. (2018) investigate the impact of customer reviews on customer evaluation under an easy-to-read review font. Hu and Chen (2016) examine review helpfulness and its relationship with other variables. Hu et al. (2017) analyze the theoretical foundation of review helpfulness and report how the interactions among three user-controllable filters together with three groups of predictors affect review helpfulness.

For other applications of hotel data to management, Cheng et al. (2010) apply context-dependent data employment analysis in measuring the performance of international hotels in Taiwan. Tseng (2009) investigates hotel service quality perceptions based on fuzzy measures of 22 evaluation criteria. Ozturk et al. (2016) test a research model that incorporates antecedents of mobile shopping loyalty in a hotel booking context. Garrido-Moreno and Padilla-Melendez (2011) examine the relationship between knowledge management and customer relationship management success using a structural equation model. Kim et al. (2016) study the relationship between the hotel industry's green practices and customer satisfaction.

The organization of this paper is as follows: Section 2 explains the mathematical setting of the overbooking model; Section 3 presents numerical examples of the overbooking strategies and the price competition with the actual online booking data of the hotels in Shinjuku ward in Tokyo; Section 4 concludes the study; and Appendices A and B provide the proofs of the theorems in Sections 2 and 3.

2. Model

In this section, we introduce an overbooking model for online booking systems with an expected sales maximization problem in which the optimal room charge and overbooking strategy are numerically obtained. We consider the expected sales maximization of a hotel for a certain room type, which considers the room charges of other hotels with similar characteristics, such as capacity, rating, and reviews, in the same area.

Additionally, we assume that the room charges of hotels are unchanged throughout the booking period, and the hotel aims to maximize its expected sales by setting its room charge and the overbooking level by taking into account room charges of the other hotels and its own capacity for the room type. Note that we adopt consistent pricing during the booking period for the transparent booking systems, where the names of the hotels to be booked are revealed to the customers. Consistency is important in a transparent pricing policy to protect loyal customers, who book in advance through the direct booking channel at a regular high price, from frequent discounting near the check-in date. This is in contrast to the dynamic pricing observed in opaque pricing systems in the U.S., where the names of the hotels are not revealed to the customers until the booking is completed (see Anderson and Xie (2012) for the dynamic pricing in opaque pricing systems).

We consider the booking behavior of online customers who randomly visit the website at a frequency following a Poisson process and book a certain room type of a hotel by choosing among a group of hotels. Let T be a check-in date and $[0,T]$ be the booking period for the check-in date, where 0 is the start date of the booking period. Then, let $J \in \mathbf{N}$ be the number of hotels in the group from which customers choose one to book, and $\{N_t\}_{0\leq t\leq T}$ be the Poisson process with intensity $\lambda > 0$, which represents the total number of rooms booked for J hotels by date t. We fix the hotel that considers the expected sales maximization as hotel i.

Let L_i^{act}, L_i^{ob} with $0 < L_i^{act} \leq L_i^{ob}$ be the actual capacity of hotel i for the room type, and the overbooking level, up to which hotel i accepts reservations for the room type in the online booking system, respectively. Let $x^{(1)}, \ldots, x^{(J)} \in (0,\infty)$ be the room charges of hotels $1,\ldots,J$ during the booking period.

2.1. The choice behavior of the customers

Firstly, let $\gamma_j \in \{0,1\}$ be the room availability of hotel j, where $\gamma_j = 0, 1$ indicates that the room is fully occupied or still available, respectively. Let $\gamma = (\gamma_1,\ldots,\gamma_J) \in \Pi_{j=1}^J\{0,1\}$ be the state of room availability of J hotels for the room type. We denote the set of availabilities of J hotels $\Pi_{j=1}^J\{0,1\}$ by Γ. We assume that $\gamma_j \equiv 1$ for all $j \neq i$, implying that hotel i maximizes its expected sales, assuming the other hotels are available for booking, and considering that hotel i does not have information about the capacity of the other hotels.

Let $\tau = (i_1, i_2, \ldots, i_{k'}) \in \mathcal{S}_{k'}$, $i_1, i_2, \ldots, i_{k'} \in \{1,\ldots,J\}$ be a selection order when k' rooms are booked in total for the J hotels in the order of $i_1, i_2, \ldots, i_{k'}$. Here, $\mathcal{S}_{k'}$ is the set of all the selection orders when k' rooms are booked in total in the group. For the selection order $\tau = (i_1,\ldots,i_{k'}) \in \mathcal{S}_{k'}$ and $l,\ 0 \leq l \leq k'$, we define $\gamma^{l,\tau} \in \Gamma$, the room availability of the J hotels after l customers booked in the order of τ, by $\gamma_j^{l,\tau} \equiv 1,\ j \neq i$, $\gamma_i^{0,\tau} = 1$, and

$$\gamma_i^{l,\tau} = \begin{cases} 1, \text{ if } \sum_{l'=1}^{l} 1_{\{i_{l'}=i\}} < L_i^{ob} \\ 0, \text{ otherwise} \end{cases}, \quad 1 \leq l \leq k',$$

which indicates that hotel i is available for choosing until the number of bookings for hotel i reaches its limit L_j^{ob}.

Moreover, let $R_T^{(j)},\ j = 1, \ldots, J$, be the number of rooms booked for hotel j by the check-in date T, which is defined by

$$R_T^{(j)} = \sum_{l'=1}^{k'} 1_{\{i_{l'}=j\}},\ j = 1, \ldots, J, \tag{1}$$

when the total number of rooms booked in the group $N_T = k'$ and the selection order $\tau = (i_1, i_2, \ldots, i_{k'})$.

Note that since the number of rooms booked for hotel i by the check-in date T does not exceed the overbooking level L_i^{ob},

$$R_T^{(i)} \leq L_i^{ob}, \tag{2}$$

and

$$\sum_{i=1}^{J} R_T^{(i)} = N_T \tag{3}$$

hold.

Furthermore, let $p_i^{(\gamma)}$ $(i = 1, \ldots, J)$ be the choice probability of hotel i, which is a function of $x^{(1)}, \ldots, x^{(J)}$ and dependent on the room availability $\gamma \in \Gamma$. For the choice probability of hotel i, $p_i^{(\gamma^{l,\tau})}$ with room availability $\gamma^{l,\tau} \in \Gamma$, we assume either of the following:

- The multinomial logit model

$$p_i^{(\gamma^{l,\tau})} = \frac{e^{-\beta x^{(i)}+\alpha_i} 1_{\{\gamma_i^{l,\tau}=1\}}}{\sum_{j=1}^{J} e^{-\beta x^{(j)}+\alpha_j} 1_{\{\gamma_j^{l,\tau}=1\}}}, \tag{4}$$

  where $\beta > 0$, $\alpha_j \in \mathbf{R}$.

- The nested logit model

$$p_i^{(\gamma^{l,\tau})} = \frac{\left(\sum_{j\in C_{k_i}} e^{\frac{-\beta x^{(j)}+\alpha_j}{\nu_{k_i}}} 1_{\{\gamma_i^{l,\tau}=1\}}\right)^{\nu_{k_i}}}{\sum_{k=1}^{n}\left(\sum_{j\in C_k} e^{\frac{-\beta x^{(j)}+\alpha_j}{\nu_k}} 1_{\{\gamma_j^{l,\tau}=1\}}\right)^{\nu_k}} \cdot \frac{e^{\frac{-\beta x^{(i)}+\alpha_i}{\nu_{k_i}}} 1_{\{\gamma_i^{l,\tau}=1\}}}{\sum_{j\in C_{k_i}} e^{\frac{-\beta x^{(j)}+\alpha_j}{\nu_{k_i}}} 1_{\{\gamma_j^{l,\tau}=1\}}} \tag{5}$$

  where $\beta > 0$, $\alpha_j \in \mathbf{R}$, $0 < \nu_k \leq 1,\ k = 1, \ldots, n$.

- The mixed logit model

$$p_i^{(\gamma^{l,\tau})} = \int_0^\infty \frac{e^{-\beta x^{(i)}+\alpha_i} 1_{\{\gamma_i^{l,\tau}=1\}}}{\sum_{j=1}^{J} e^{-\beta x^{(j)}+\alpha_j} 1_{\{\gamma_j^{l,\tau}=1\}}} h(\beta) d\beta, \tag{6}$$

  where $\alpha_j \in \mathbf{R}$, $0 \leq h(\beta)$, $\int_0^\infty h(\beta)d\beta = 1$.

We remark that in the theory of random utility (e.g., See Train (2009)), the choice probabilities $p_i^{(\gamma^{l,\tau})}, i = 1, \ldots, J$ in the multinomial logit model (4) correspond to the situation in which customers choose the hotel with the highest utility, and the customers' utility of hotel j $(j = 1, \ldots, J)$ is given by $U_j = -\beta x^{(j)} + \alpha_j + \epsilon_j$. Here, ϵ_j is a random variable where $\{\epsilon_j\}_{j=1,\ldots,J}$ are independent and identically distributed (IID) random variables following an extreme value distribution and the deterministic part of the random utility, $-\beta x^{(j)}+\alpha_j$, is a decreasing function on the room charge $x^{(j)}$, which implies that the higher the room charge is, the lower the utility is. The other characteristics of hotel j, which do not change over time such as reputation and grade, are reflected in α_j. The indicator function in (4) implies that if $\gamma_i^{l,\tau} = 0$ (i.e., hotel i is fully booked and not available for booking), then hotel i is excluded from the group for the choice.

Further, the nested logit model in (5) is interpreted in terms of the random utility theory as follows. The J hotels are divided into n non-overlapping groups of the same kind $\{C_k\}_{k=1,\dots,n}$, and the dissimilarities within the small groups are represented by $0 \le \nu_k \le 1,\ k = 1, \dots, n$. The first term on the right-hand side of (5) represents the probability for which group C_{k_i} is selected among n small groups, and the second term corresponds to the conditional probability for which hotel i is selected within the group C_{k_i}. The model agrees with the multinomial logit model when the dissimilarity parameters are $\nu_k = 1,\ k = 1, \dots, n$, meaning that the hotels in the small groups are all dissimilar.

The mixed logit model in (6) is another generalization of the multinomial logit model in (4), in which the sensitivity of the utilities to change in the room charge β takes different values among customers with the distribution $h(\beta)$. We use this mixed logit model in (6) in the numerical examples and explain it in detail in Section 3.2.

We assume that the customers' choice behavior is independent of the total number of bookings for the group by the check-in date T. Namely, the conditional probability for the choice order $\tau = (i_1, i_2, \dots, i_{k'})$ when $N_T = k'$ is given by

$$\mathbf{P}\left(\tau = (i_1, i_2 \dots, i_{k'}) | N_T = k'\right) = p_{i_1}^{(\gamma^{0,\tau})} p_{i_2}^{(\gamma^{1,\tau})} \cdots p_{i_{k'}}^{(\gamma^{k'-1,\tau})}, \tag{7}$$

and the probability for the choice order $\tau = (i_1, i_2, \dots, i_{k'})$ is

$$\mathbf{P}\left(\tau = (i_1, i_2 \dots, i_{k'})\right) = p_{i_1}^{(\gamma^{0,\tau})} p_{i_2}^{(\gamma^{1,\tau})} \cdots p_{i_{k'}}^{(\gamma^{k'-1,\tau})} \frac{(\lambda T)^{k'}}{k'!} e^{-\lambda T}. \tag{8}$$

Note that since the choice probability $p_i^{(\gamma)}$ includes the room charges $x^{(j)},\ j = 1, \dots, J$ and the room availability γ, $\mathbf{P}\left(\tau = (i_1, i_2 \dots, i_{k'}) | N_T = k'\right)$ in (7) is a function of $x^{(j)},\ j = 1, \dots, J$ and L_i^{ob}.

Remark 1. *We assume that λ, the intensity of N_t which represents the total number of bookings for the J hotels until t, is a constant and does not depend on $x^{(1)}, \dots, x^{(J)}$, the room charges of the hotels. In other words, there is no increase or decrease in the total number of customers visiting the website to choose one from the group of hotels when the overall level of the room charges decreases or increases. This corresponds to the situation in which the customers visiting the booking website aim to choose a hotel from a selection of hotels with similar prices in the same area. Extending the model to apply it to the case with the intensity of λ dependent on the room charges will be one of our research topics.*

2.2. Maximization of the expected sales

Next, we consider the expected sales of the hotel i.

Firstly, let r_i be the last-minute cancellation rate of hotel i, which is a $\mathbf{R}$-valued random variable. We assume that the cancellation rate r_i has a discrete probability distribution with a support $[0, \bar{r}_i]$, $0 < \bar{r}_i < 1$ and is independent of the total number of rooms booked for the group N_T and that for hotel i, $R_T^{(i)}$. Then, the number of rooms booked on the check-in date T after the last-minute cancellations is

$$(1 - r_i) R_T^{(i)}. \tag{9}$$

Secondly, let c_i be the over-sale cost per room of hotel i, which is a positive constant. Namely, hotel i incurs the cost when the number of rooms booked after the cancellations exceeds the actual room capacity L_i^{act}.

Specifically, i) if the number of bookings after last-minute cancellations does not exceed the actual capacity,

$$(1 - r_i) R_T^{(i)} \le L_i^{act}, \tag{10}$$

there is no over-sale cost.

ii) If the number of bookings after last-minute cancellations exceeds the actual capacity, namely,

$$(1 - r_i)R_T^{(i)} > L_i^{act}, \tag{11}$$

then hotel i incurs

$$\left((1 - r_i)R_T^{(i)} - L_i^{act}\right) c_i \tag{12}$$

as the over-sale cost. By combining the two cases, the sales for hotel i including the over-sale cost are

$$\min\left(L_i^{act}, (1 - r_i)R_T^{(i)}\right) x^{(i)} - \max\left((1 - r_i)R_T^{(i)} - L_i^{act}, 0\right) c_i. \tag{13}$$

Here, we remark that when hotel i has spare rooms in higher grade room types and they can be used for upgrading, the over-sale cost per room c_i is low. Otherwise, the hotel must displace the guests to other hotels by paying for the room charges and may offer extra compensations, and the over-sale cost per room c_i is high.

Finally, the optimization problem of hotel i is defined as follows. Hotel i aims to maximize the expectation of the sales including the over-sale cost by choosing the room charge $x^{(i)} \in (0, \infty)$ and the overbooking level $L_i^{ob} \in [L_i^{act}, \infty)$, given the room charges of the other hotels $\{x^{(j)}\}_{j \neq i}$. Namely, we find the pair of $(x^{(i)*}, L_i^{ob*}) \in (0, \infty) \times [L_i^{act}, \infty)$ that attains

$$\max_{(x^{(i)}, L_i^{ob}) \in (0,\infty) \times [L_i^{act}, \infty)} \mathbf{E}\left[\min\left(L_i^{act}, (1 - r_i)R_T^{(i)}\right) x^{(i)}\right] - \mathbf{E}\left[\max\left((1 - r_i)R_T^{(i)} - L_i^{act}, 0\right) c_i\right]. \tag{14}$$

Remark 2. *We remark that if r_i is a constant, the optimal overbooking level is $L_i^{ob} = L_i^{act}/(1-r_i)$ since as long as $L_i^{ob} \leq L_i^{act}/(1 - r_i)$, overselling cannot not happen, and it is better for hotel i to accept as many bookings as possible. For $L_i^{ob} > L_i^{act}/(1 - r_i)$, overselling occurs and the hotel incurs more over-sale cost as it raises the overbooking level. Hence, it is optimal for hotel i to set L_i^{ob} at this level if the cancellation rate is a constant. The distribution of the cancellation rate r_i as well as the over-sale cost per room c_i can be provided exogenously based on hotel i's own empirical data.*

The next theorem guarantees the existence of the optimal room charge and overbooking level for the maximization problem of hotel i given $\{x^{(j)}\}_{j \neq i, j \in \{1,\ldots,J\}}$ and L_i^{act}, the room charges of the other hotels and the actual capacity of hotel i. The theorem indicates that we can find optimal $(x^{(i)*}, L_i^{ob*})$ numerically by taking a large enough bounded domain, dividing it into a mesh, calculating the values of the objective function at all girds, and taking the maximum.

Theorem 1. *Suppose that for all $\gamma \in \Gamma$ and $j = 1, \ldots, J$, $p_j^{(\gamma)}$ is given either by the multinomial logit model in (4), the nested logit model in (5), or the mixed logit model in (6) with*

$$\int_0^\infty \frac{1}{\beta} h(\beta) d\beta < \infty. \tag{15}$$

Then, there exists $(x_*^{(i)}, L_{i*}^{ob})$ in $(0, \infty) \times [L_i^{act}, \infty)$, which attains the maximum of

$$\mathbf{E}\left[\min\left(L_i^{act}, (1 - r_i)R_T^{(i)}\right) x^{(i)}\right] - \mathbf{E}\left[\max\left((1 - r_i)R_T^{(i)} - L_i^{act}, 0\right) c_i\right] \tag{16}$$

Proof. See Appendix A.

Moreover, the proof in Appendix A shows that the objective function decreases to 0 when $x^{(i)}$ goes to ∞ because of the exponential decrease in the choice probability of hotel i. It is also shown that when the overbooking level L_i^{ob} goes to ∞, although hotel i can accept as many rooms as possible, the increase in the expected over-sale cost (the second term of the objective function) outweighs the increase in the expected sales (the first term), and the net amount is decreasing after some overbooking level.

3. Numerical examples

This section presents numerical examples of the optimal room charge and overbooking level of a hotel by using actual online booking data of nonsmoking standard twin rooms for two major luxury hotels in Shinjuku ward in Tokyo, which are crawled from a Japanese online booking website. We also investigate the maximization problem for four different scenarios of over-sale cost per room and distribution of the cancellation rate of the hotel.

3.1. Data set

The original dataset includes prices of the accommodation plans and available numbers of accommodation plans for booking for the nonsmoking standard twin rooms at Hotels A and B in Shinjuku ward in Tokyo. The check-in dates and the booking dates of the accommodation plans in the dataset range from March 1st to April 30th and 25th, 2017, respectively. We collected the data from a major Japanese hotel-booking website by crawling the publicly available online booking information.

The crawling of the online booking data is performed using the following procedure.

- First, we access the website using 20 virtual servers, and specify the necessary inputs for the search engine in the online booking website, such as the area and the district of the hotels (Shinjuku ward in Tokyo for example) and the number of guests (1 or 2).
- Then, we search for available hotels, accommodation plans, the room charges, and number of available plans, and we save the results in HTML format.
- We iterate the procedure automatically by running a program by changing the check-in date from the booking date to 60 days after the booking date.
- We repeat running the program for the dates of March 1st to April 25th, 2017.
- We implement the crawling at a frequency so as not to disturb the online websites' business.

For the check-in dates, we use the two-month data ranging from March 1st to April 30th, 2017 for the check-in dates. This is because the hotel demand has a seasonality affected by tourism demand. Thus, before starting the estimation of parameters in the revenue management model, we must investigate the seasonality and carefully select an appropriate period with stable demand, which turns out to be a relatively short period of one to two months or shorter. In addition, when we use publicly available data by crawling data from online booking systems in research, we must work on the analysis within our limitation of the data available. Unless we have a special relationship with a hotel, it is difficult to collect data over a long period. As for the seasonality of tourism and hotel demand, see Goh and Law (2002), Duro (2016), and Sierag et al. (2017), for example.

For studies on revenue management that consider the seasonality and publicly available data crawled from online booking websites, Abrate and Viglia (2016) conducted an empirical analysis of the factors affecting the setting of the room charges of hotels in Milan by using data from 16 days in October 2012. Furthermore, Viglia et al. (2016) analyzed the impact of online reviews on occupancy rates of hotels in Rome using data from the month of November 2014.

March and April in Japan form the busy tourist season due to spring vacation, graduation trips for students, and cherry blossom viewing. We choose this season for its stable demand for hotels, following the quiet winter season with fewer tourists and preceding the Golden Week holiday season in May, when the number of tourists surges and the room charges soar.

Hotels A and B are competing in the same area and have similar characteristics as they are both luxury hotels with large capacities, high ratings, and good customer reviews. Table 1 describes the basic information of Hotels A and B (e.g., total number of guest rooms for all room types, ratings, and customer reviews).

	Number of rooms	Booking.com rating	Michelin guide 2011 rating	Booking.com review	Rakuten review	Jalan review
Hotel A	806	5 stars	4 pavilions	8.5/10.0	3.88/5.00	3.9/5.0
Hotel B	744	4 stars	3 pavilions	8.6/10.0	4.17/5.00	4.5/5.0

Table 1: Summary of Hotels A and B

Secondly, we note that the original dataset does not explicitly contain the information on the room charges or the numbers of rooms sold for this room type, while it includes the prices and the number of plans available for booking for the multiple accommodation plans linked to this room type. Thus, to use this model, we need to estimate the information from the original dataset.

Taking these points into consideration, we first define the representative room charge of this room type of a hotel for each check-in date as follows. For each check-in date and accommodation plan, we calculate the average price of the plan over the corresponding booking dates. Then, we take the minimum of the average prices over the accommodation plans for each check-in date and designate it as the representative room charge of the room type for the check-in date. We use the representative room charges as the room charges in the model.

Next, we define the number of rooms booked for this room type of a hotel on a booking date for a check-in date as follows. We first calculate the change of the available number of plans for an accommodation plan from the previous to the current booking date. If the change is negative, we regard this as the number of rooms booked on the booking date, and if this is positive, we consider that this number of new rooms were supplied by the hotel to the booking systems. Note that the number of plans available for booking of an accommodation plan changes in conjunction with all the other accommodation plans linked to the same room type, and hence the numbers are the same among all the accommodation plans.

Finally, Table 2 summarizes the information on the dataset including the representative room charges and the numbers of rooms booked for the nonsmoking standard twin rooms of Hotels A and B, which will be used in the estimation of the model parameters in Section 3.2. Here, "Number of check-in days" in Table 2 indicates the number of check-in dates used in the estimation, which is 61 days from March 1st to April 30th. The "number of sold-out check-in days" expresses the number of check-in dates when the room type analyzed has been fully occupied by the check-in date or by April 25th for the check-in dates from April 26th to 30th. This dataset used for the analysis in Section 3.2 is available at Mendeley Data (http://dx.doi.org/10.17632/9tkynnhj23.1).

	Hotel A	Hotel B
Number of check-in days	61	61
Number of check-in days, a day before a holiday	9	9
Number of sold out check-in days	30	35
Number of sold out check-in days, a day before a holiday	5	7
Number of booking days	1,876	1,876
Number of available booking days	976	682
Maximum representative room charge	62,911	59,400
Minimum representative room charge	26,922	20,925
Average representative room charge	40,799	31,917
Maximum representative room charge, a day before a holiday	55,719	59,400
Minimum representative room charge, a day before a holiday	36,355	30,645
Average representative room charge, a day before a holiday	43,447	33,950
Maximum representative room charge, a weekday	62,911	50,202
Minimum representative room charge, a weekday	26,922	20,925
Average representative room charge, a weekday	40,367	31,721
Maximum number of booked rooms per check-in day	20	21
Minimum number of booked rooms per check-in day	0	0
Average number of booked rooms per check-in day	3.97	4.93
Maximum number of booked rooms per check-in day, a day before a holiday	10	10
Minimum number of booked rooms per check-in day, a day before a holiday	0	0
Average number of booked rooms per check-in day, a day before a holiday	3.33	3.00
Maximum number of booked rooms per check-in day, a weekday	20	21
Minimum number of booked rooms per check-in day, a weekday	0	0
Average number of booked rooms per check-in day, a weekday	4.08	5.27

Table 2: Summary of the dataset used for the estimation

3.2. Estimation

Firstly, in the following numerical examples, we assume the mixed-logit model (6) with the log-normal distribution for β, that is, $\beta = e^X$ where X is a random variable following the normal distribution with the mean $\mu \in \mathbf{R}$ and the standard deviation $\sigma > 0$, for $p_j^{(\gamma)}$, the choice probability of hotel j, $j = 1, \ldots, J$, with the room availability $\gamma \in \Gamma$. We remark that the integrability condition (15) on β of the mixed logit model (6) in Theorem 1 holds for this lognormal distribution.

In addition, we assume that the intercept α_j, $j = 1, \ldots, J$ in (6) takes the form $\alpha_j = \delta_j y + \bar{\alpha}_j$, where y is a dummy variable taking a value 0 or 1. Namely, $y = 0$ if the check-in date is a weekday and $y = 1$ if it is a day before a holiday. Particularly, we label Hotels A and B as hotels 1 and 2, respectively, and consider the case of $J = 2$.

Then,

$$
\begin{aligned}
p_j^{(\gamma)} &= \int_0^\infty \frac{e^{-\beta x^{(j)}+\delta_j y+\bar{\alpha}_j} 1_{\{\gamma_j=1\}}}{\sum_{j'=1}^2 e^{-\beta x^{(j')}+\delta_{j'} y+\bar{\alpha}_{j'}} 1_{\{\gamma_{j'}=1\}}} h(\beta) d\beta, \\
&= \int_0^\infty \frac{e^{-e^{(\mu+\sigma z)} x^{(j)}+\delta_j y+\bar{\alpha}_j} 1_{\{\gamma_j=1\}}}{\sum_{j'=1}^2 e^{-e^{(\mu+\sigma z)} x^{(j')}+\delta_{j'} y+\bar{\alpha}_{j'}} 1_{\{\gamma_{j'}=1\}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz
\end{aligned} \tag{17}
$$

for $j = 1, 2$. Noting that $p_j^{(\gamma)}$ depends on the differences $\bar{\alpha}_2 - \bar{\alpha}_1$ and $\delta_2 - \delta_1$ for $\bar{\alpha}_j$ and δ_j, we estimate the parameters $\mu, \sigma, \bar{\alpha}_2$, and δ_2 by fixing $\bar{\alpha}_1 = \delta_1 = 0$.

Specifically, we can interpret the mixed logit model in (6), a generalization of the multinomial logit model incorporating differences of the price sensitivity β among the customers, in the framework of the random utility theory (e.g., Train (2009)) as follows. Suppose that the utility functions on Hotels A and B of person $m \in M$, $U_{1,m}$ and $U_{2,m}$, are $U_{j,m} = -\beta_m x^{(j)} + \delta_j y + \bar{\alpha}_j + \epsilon_{j,m}$, where $\epsilon_{j,m}$ is the random term of the utility and $\{\epsilon_{j,m}\}_{j=1,2}$ are IID random variables following an extreme value distribution. β_m is the price sensitivity parameter of person m. Person m chooses Hotel A if $U_{1,m} > U_{2,m}$, and Hotel B otherwise. Then, it follows that the choice probabilities of Hotels A and B, $p_{1,m}^{(\gamma)} = \mathbf{E}[U_{1,m} > U_{2,m}]$ and $p_{2,m}^{(\gamma)} = \mathbf{E}[U_{2,m} > U_{1,m}]$, by person m under the room availability γ are given by

$$
p_{j,m}^{(\gamma)} = \frac{e^{-\beta_m x^{(j)}+\delta_j y+\bar{\alpha}_j} 1_{\{\gamma_j=1\}}}{\sum_{j'=1}^2 e^{-\beta_m x^{(j')}+\delta_{j'} y+\bar{\alpha}_{j'}} 1_{\{\gamma_{j'}=1\}}}. \tag{18}
$$

Assuming that β_m is a sample of a lognormal random variable $e^X, X \sim N(\mu, \sigma)$ and taking average of the choice probabilities $p_i^{(\gamma)}$ $(i = 1, 2)$ over all persons M, we have

$$
\begin{aligned}
p_j^{(\gamma)} &= \int_0^\infty \frac{e^{-\beta x^{(j)}+\delta_j y+\bar{\alpha}_j} 1_{\{\gamma_j=1\}}}{\sum_{j'=1}^2 e^{-\beta x^{(j')}+\delta_{j'} y+\bar{\alpha}_{j'}} 1_{\{\gamma_{j'}=1\}}} h(\beta) d\beta \\
&= \int_0^\infty \frac{e^{-e^{(\mu+\sigma z)} x^{(j)}+\delta_j y+\bar{\alpha}_j} 1_{\{\gamma_j=1\}}}{\sum_{j'=1}^2 e^{-e^{(\mu+\sigma z)} x^{(j')}+\delta_{j'} y+\bar{\alpha}_{j'}} 1_{\{\gamma_{j'}=1\}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz.
\end{aligned} \tag{19}
$$

Finally, we estimate the parameters $\mu, \sigma, \bar{\alpha}_2$, and δ_2 of the mixed-logit model (19) by the maximum simulated likelihood method (See Train (2009) for instance), and Table 3 shows the estimation results.

	Estimate	Std. Error	t-value	p-value
μ	-8.161672	0.087479	-93.2985	< 2.2e-16 ***
σ	1.053053	0.243032	4.333	1.471e-05 ***
$\bar{\alpha}_2$	-2.094104	0.27978	-7.4848	7.172e-14 ***
δ_2	0.849365	0.604821	1.4043	0.1602

Note: * - marginal significant at 0.10 level; ** - significant at 0.05 level; *** - significant at 0.01 level.

Table 3: Estimation results of the mixed-logit model.

In addition to the parameters related on the choice behaviors of the customers, we set the other parameters as $\lambda = 2.1429, T = 14, L_1^{act} = 20, x^{(2)} = 42{,}292$, which correspond to the booking data of the check-in date April 10th, 2017 when the nonsmoking standard twin rooms of Hotel A are fully booked after 20 rooms were sold and 10 rooms are sold for Hotel B with the representative room charges JPY 46,680 and JPY 42,292 for Hotels A and B, respectively. λ is also estimated by the maximum likelihood method.

3.3. Results of the expected sales maximization problem

Next, we calculate the optimal room charge and overbooking level of Hotel A in four scenarios of the cancellation rate and the over-sale cost per room. We set the probability distributions of the cancellation rates and the levels of the over-sale cost per room as follows: $\mathbf{P}(r_1^H = 70\%) = \mathbf{P}(r_1^H = 50\%) = \mathbf{P}(r_1^H = 30\%) = \frac{1}{3}$

for the high cancellation rate r_1^H , $\mathbf{P}(r_1^L = 0\%) = \mathbf{P}(r_1^L = 10\%) = \mathbf{P}(r_1^L = 20\%) = \frac{1}{3}$ for the low cancellation rate r_1^L, the high over-sale cost per room $c_1^H = 100{,}000$, and the low over-sale cost per room $c_1^L = 100$, where the costs are expressed in JPY.

Firstly, Figures 1-4 show the graphs of the objective function on the room charge and overbooking level in four cases, the high/low cancellation rate and the high/low over-sale cost per room. The maximum points are numerically obtained by searching for the maximum point of the objective function in a bounded domain $(0, 75{,}000] \times [20, 100] \subset (0, \infty) \times [L_1^{act}, \infty)$, which is broad enough as shown in the four graphs below.

1. The low cancellation rate and the low over-sale cost per room: JPY 824,200 when $(x_*^{(1)}, L_{1*}^{ob}) = (43{,}000, 25)$.

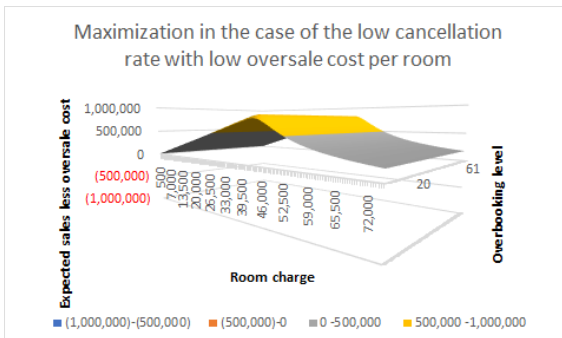

Figure 1: The expected sales in the case of the low cancellation rate and the low over-sale cost per room

2. The low cancellation rate and the high over-sale cost per room: JPY 763,244 when $(x_*^{(1)}, L_{1*}^{ob}) = (43{,}500, 20)$.

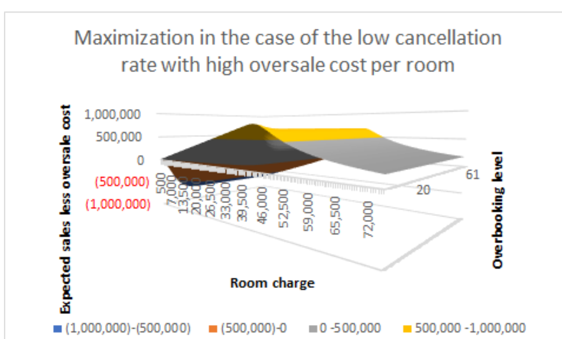

Figure 2: The expected sales in the case of the low cancellation rate and the high over-sale cost per room

3. The high cancellation rate and the low over-sale cost per room: JPY 555,221 when $(x_*^{(1)}, L_{1*}^{ob}) = (41{,}500, 51)$.

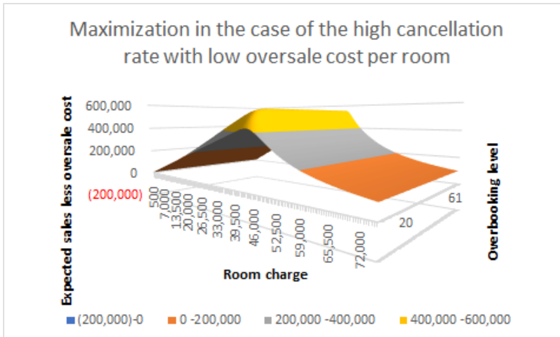

Figure 3: The expected sales in the case of the high cancellation rate and the low over-sale cost per room

4. The high cancellation rate and the high over-sale cost per room: JPY 536,454 when $(x_*^{(1)}, L_{1*}^{ob}) = (42{,}000, 29)$.

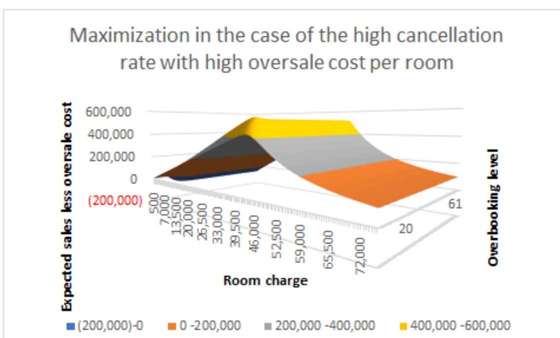

Figure 4: The expected sales in the case of the high cancellation rate and the high over-sale cost per room

In detail, Figure 5 illustrates the expected sales for different room charges in the case of the high cancellation rate r_1^H and the high over-sale cost per room c_1^H in the third case above, when the overbooking level is set at the optimal level, $L_{1*}^{ob} = 29$.

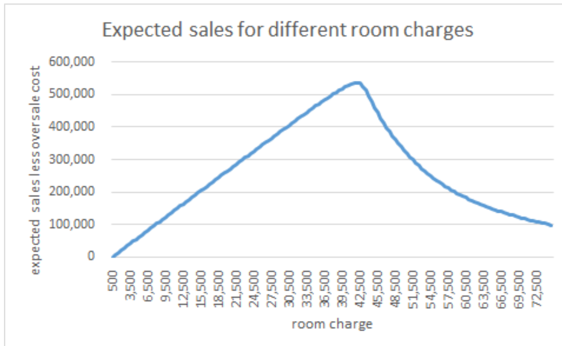

Figure 5: The expected sales for different room charges

Then, we observe that the expected sales have a peak of JPY 536,454 at a room charge of JPY 42,000. At first, as the room charge increases, the expected sales also increase because of the higher unit price. After the optimal room charge, the expected sales decline, since the decrease in choice probability outweighs the increase in unit price.

Moreover, Figures 6 and 7 show the optimal expected sales for different overbooking levels in the four cases, where the expected sales are maximized over the room charge. Then, we observe the following.

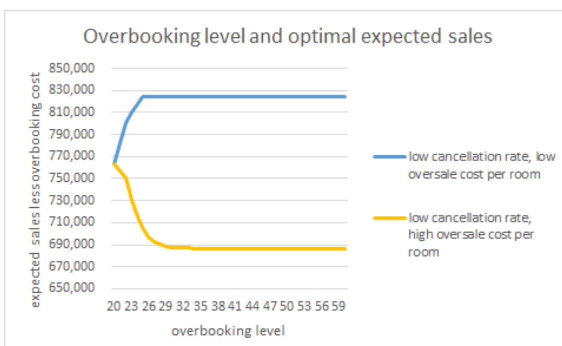

Figure 6: The optimal expected sales less over-sale cost for different overbooking levels, low cancellation rate cases

- Firstly, with the low cancellation rate, in the case of the low over-sale cost per room, as the blue line indicates, overbooking compensates the sales loss by the last-minute cancellations and this compensation outweighs the cost by the over-sale up to 25 rooms of the overbooking level. After that, there is no more sales compensation by raising the overbooking level, and the expect sales only slightly decrease because of the over-sale cost. In the case where the over-sale cost per room is high, as the yellow line illustrates, the expected sales only decrease. Due to the low cancellation rate, there is not large sales loss by the cancellations, and because the over-sale cost per room is high, over-sale cost outweighs the sales compensation from 20 rooms of the overbooking level.

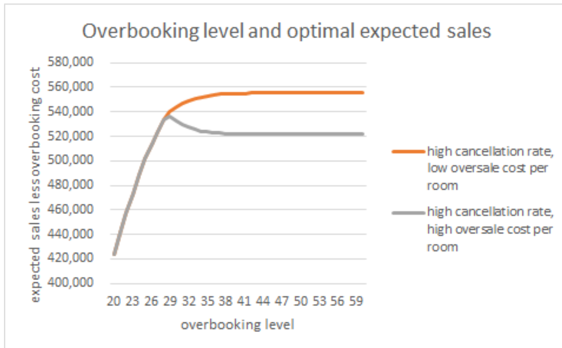

Figure 7: The optimal expected sales less over-sale cost for different overbooking levels, high cancellation rate cases

- Secondly, with the high cancellation rate, as the orange and gray lines indicate, in the case of the low over-sale cost, up to 51 rooms of the overbooking level, the compensation effect outweighs the over-sale cost and the orange line is almost flat due to the low over-sale cost per room and the small chance for the rooms to be booked up to this level. However, when the over-sale cost per room is high, as the gray line shows, after 29 rooms of the overbooking level, overbooking results in high over-sale cost and the expected sales start to decline.

As we have observed, the results in the four cases not only describe the trade-offs on the room charge and the overbooking level but also provides concrete quantitative levels of the optimal overbooking and room charge.

3.4. Equilibrium room charges

This section presents equilibrium room charges of Hotels A and B by investigating how the room charges shift and where they converge if each hotel revises its room charge in response to the optimization of the counterpart.

Firstly, in a particular case of the multinomial logit model in (4) with $\gamma \equiv (1,1)$ for the choice probability, implying that there is no restriction on the room availability for the hotels, the next theorem indicates that a unique Nash equilibrium exists and the room charges converge to the equilibrium levels by the iterative maximizations of the hotels.

Theorem 2. *Suppose that the choice probabilities of hotels 1 and 2 are given by the multinomial logit model*

$$\begin{aligned} p_1 &= \frac{1}{1+e^{\beta(x^{(1)}-x^{(2)})+\alpha_2-\alpha_1}}, \\ p_2 &= \frac{1}{1+e^{\beta(x^{(2)}-x^{(1)})+\alpha_1-\alpha_2}}, \end{aligned} \tag{20}$$

and hotels 1 and 2 maximize their expected sales $\mathbf{E}[x^{(1)}R_T^{(1)}]$ *and* $\mathbf{E}[x^{(2)}R_T^{(2)}]$*, respectively, given the room charge of the counterpart. Then, a unique Nash equilibrium* $(x_*^{(1)}, x_*^{(2)}) \in (\frac{1}{\beta}, \infty)^2$ *exists. Moreover, if hotels 1 and 2 maximize their expected sales iteratively starting from the room charges* $(x_0^{(1)}, x_0^{(2)}) \in (0,\infty)^2$*, the room charges converge to the Nash equilibrium.*

Here, we call (x_1^*, x_2^*) *a Nash equilibrium for the room charges if they satisfy*

$$x_*^{(1)} = argmax_{x^{(1)}\in(0,\infty)} f(x^{(1)}|x_*^{(2)}), \tag{21}$$

$$x_*^{(2)} = argmax_{x^{(2)}\in(0,\infty)} g(x^{(2)}|x_*^{(1)}), \tag{22}$$

where $f(x^{(1)}|x^{(2)})$ and $g(x^{(2)}|x^{(1)})$ are the expected sales of hotels 1 and 2 given the room charge of the counterpart.

Proof. See Appendix B.

Taking these into consideration, we numerically investigate the equilibrium for the room charges by calculating the transition of the optimal room charges in the case of the high cancellation rate and the over-sale cost per room in Section 3.3, which is a generalized case of Theorem 2, in the following manner. Firstly, we calculate the optimal room charge and overbooking level of Hotel A maximizing its expected sales given the initial room charge of Hotel B. Then, we calculate the optimal room charge of Hotel B given the optimal room charge of Hotel A. We iterate this optimization process 10 times in total.

Firstly, Figure 8 illustrates the transition of the optimal room charges of Hotels A and B by the iterative optimizations. Starting from the initial room charge of JPY 46,680 and JPY 42,292 for Hotels A and B, Hotel A optimizes its room charge given the initial room charge of Hotel B, and Hotel B optimizes given the optimized room charge of Hotel A. The room charges obtained, JPY 42,000 and JPY 30,000 for Hotels A and B, respectively, are shown as the optimal room charges in the first iteration. In the second iteration, Hotel A optimizes its room charge given the optimized room charge of Hotel B after the first iteration, and then Hotel B optimizes given the optimized room charge of Hotel A. They iterate this process 10 times in total. We observe that the optimal room charges decline by the iterative optimizations and converge to the room charges, JPY 10,500 and JPY 6,500, which are at lower levels than the original room charges.

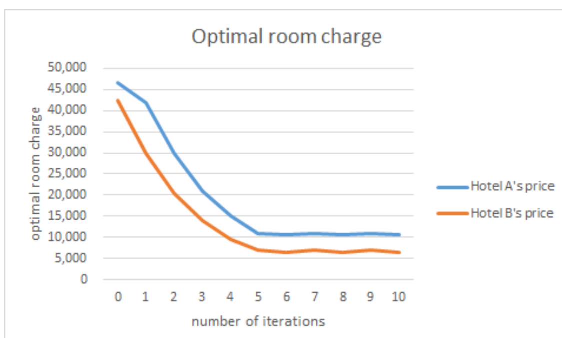

Figure 8: Convergence of the optimal room charges

Secondly, Figure 9 describes the corresponding expected sales. In accordance with the declines in the room charges, the expected sales also decrease for both hotels. Specifically, the expected sales start from the initial room charges of JPY 407,992 and JPY 528,751 and then increase to JPY 536,454 and JPY 620,827 in the first optimization for Hotels A and B, respectively. However, thereafter they decrease to JPY 99,654 and JPY 77,084, respectively.

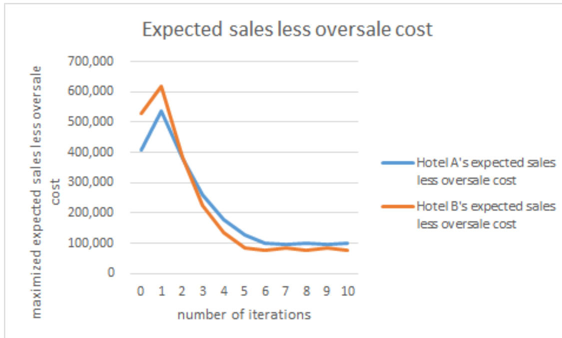

Figure 9: Transition of the expected sales by the iterative optimizations

Finally, to better understand how the optimal room charges converge, Figure 10 displays the best responses of the hotels, that is, the correspondence between the optimal room charge and the given room charge of the counterpart. Starting from the first optimization of Hotel A, which is the optimal room charge JPY 42,000 in response to the initial room charge JPY 42,292 of Hotel B, the room charges converge to the crossing point of the two graphs, JPY 10,500 and JPY 6,500 for Hotels A and B, respectively.

Figure 10: Optimal room charges of Hotels A and B in response to the counterpart's room charge

Particularly, these results imply that the revenues of the hotels may decrease once they are in a price competition repetitively optimizing their overbooking levels and room charges to maximize their expected sales.

4. Discussion

In this study, we have presented an application of online hotel booking data, which are big data publicly available on online booking websites, to a quantitative revenue management model for online hotel booking systems. Additionally, we have proposed a methodology to calculate the room charge and overbooking level that maximizes a hotel's expected sales with this model. The model incorporates room charges of rival hotels and their booking status estimated from the crawled online booking data of hotels.

4.1. Contributions to existing theory and literature

Existing literatures, such as Rothstein (1974), Liberman and Yechiali (1978), Gallego and Ryzin (1994), Koide and Ishii (2005), and Sierag et al. (2015), propose revenue management models for a single hotel, which calculate either the room charge or the overbooking level that maximizes the expected sales based on their own booking data. In contrast, Rothstein (1974), Liberman and Yechiali (1978), Koide and Ishii (2005), Gallego and Ryzin (1994), and Sierag et al. (2015) deal with maximization of expected revenue on a single variable—the overbooking level for Rothstein (1974), Liberman and Yechiali (1978), Koide and Ishii (2005); the room charge for Gallego and Ryzin (1994); or a set of products to sell for Sierag et al. (2015). Our study deals with the revenue maximization on the two variables, the room charge and overbooking level, simultaneously. Moreover, our study deals with revenue management under overbooking in online booking systems—where the customers choose a hotel from a group taking the room charges of the rival hotels into account—and investigates the equilibrium room charges. As for the practices in hotel revenue management, Gehrels and Blanar (2013) analyze the cases of a Hilton chain hotel, Ferguson and Smith (2014) study the history of the roles of the hotel revenue managers, and El Haddad (2015) and Ivanov and Ayas (2017) investigate the practices of the revenue managers. As those studies explain, revenue management is generally done by their rule of thumb or by the revenue management system they introduce from vendors. By adopting our methodology, the hotel revenue managers can observe expected sales which account for room charges of rival hotels in a transparent manner and that also can quantitatively monitor the changes in the expected sales for different overbooking levels.

To the best of our knowledge, this is the first attempt to calculate the optimal overbooking level and room charge for a certain room type of a hotel with the actual data crawled from an online booking website. In addition, with this model, booking patterns—including arrival rates and a price sensitivity of the customers to the choice probabilities—are estimated from the big data in the online booking system. In combination with the empirical data of the cancellation rate and the over-sale cost per room hotels own, hotels can quantitatively obtain the optimal room charge and overbooking level that maximize their expected sales. In other words, our methodology effectively utilizes further information on rival hotels by crawling publicly available big data from an online booking website and conducts expected sales maximization of the hotel by constructing the new quantitative revenue management model, which elaborates the existing revenue management models.

4.2. Implications for practice

This study has the following practical implications. With the methodology, a hotel can accumulate information on rival hotels' accommodation plans and prices along with booking status by crawling online booking websites. Combining the crawled data with the hotel's own data, which include cancellation status and over-sale costs, the hotel can calculate the room charge and overbooking level that can maximize the expected sales.

Hotels usually set room charges and overbooking levels based on their own data or by revenue management systems, which they introduce from a vendor and whose algorithms they do not know. Specifically, the two hotels in the numerical examples in Section 3 have a revenue management team and own a revenue management system. They set room charges and overbook rooms by checking the status of their inventories and how their rooms are booked. Both hotels introduce a revenue management system produced by a Japanese revenue management system vendor. The revenue managers use the system to determine the room charges and the overbooking levels by accounting for the sales predictions indicated by the system. In the revenue management system of the vendor, the revenue managers can see the sales prediction for different room charges and monitor the room charges of the rival hotels listed in online booking websites. However, in the system, both the historical booking status and the room charges of the rival hotels are not taken into consideration in the sales prediction. The room charges of the rival hotels are shown only as reference prices. Moreover, the system does not suggest any quantitative overbooking levels.

If they introduce our methodology, the hotel revenue managers can observe the expected sales that account for room charges of the rival hotels and can also monitor the changes in the expected sales for different overbooking levels. Namely, this use of the revenue management model offers more sophisticated insight into room charges, overbooking levels, and expected sales. This is expected to contribute to the hotel's revenue growth.

Remark 3. *We find that even if the hotels change their pricing algorithm, it would not affect the way we estimate the relation between the room charges and the booking probabilities of the hotels in our methodology since we estimate it by the mixed logit model, given the numbers of rooms booked and the room charges set by the hotels. Naturally, if the hotels change their pricing algorithm, the room charges they set and how the rooms are booked would change. However, since we estimate the model parameters given the data after the change in the algorithm, the hotels can improve their pricing by considering factors not incorporated into their algorithm.*

4.2.1. Implications for hotels

In addition to the implications mentioned in Section 4.1, hotels can take some internal measures to run the revenue management model effectively.

Firstly, the revenue management department can provide detailed information on how to set room charges by the revenue management model to the management and the sales & marketing department. The revenue management department can also make daily reports on the price analysis as well as the sales of the hotel and the rival hotels, and the total demand from the online customers in the area. Then, it can explain those to the management as well as the sales & marketing department to search for business opportunities. Additionally, by analyzing the data on room charges and status of booking of other areas obtained by crawling, hotels can identify the revenue and the competitiveness in the areas and utilize the data to search for new business opportunities.

Secondly, by giving an overview of the revenue management that hotels adopt to set room charges and overbooking levels, the hotels can make their employees feel secure about the organization they work for, allowing them to focus on their roles.

Thirdly, since good customers' reviews enhance hotels' probability of being booked, it is important for hotels to obtain better online reviews in cooperation with relevant internal departments. For example, by internally sharing information on customers who booked through an online booking system with the operations department, which directly faces the customers during their stay, hotels can provide the customers with special treatment that makes them comfortable and can lead to garnering good reviews and gaining loyalty from the customers. For studies on the relations between online reviews of hotels and revenues, ratings, and customer relationship management, see Neirotti et al. (2016), Xu et al. (2017), Kim et al. (2016), Hu et al. (2016), and Chang et al. (2017).

Furthermore, hotels can analyze reviews that customers previously posted for other hotels to know in advance what they care about and what makes them feel uncomfortable and, in this way, a hotel can build and maintain customer relationships. Additionally, the revenue management department should work with the employees for smooth operations in case they get feedbacks from the customers on the room charges or in case the hotels displace the customers due to over-sale and try to minimize the damage on the customer relationship management.

Finally, to gain customer loyalty, the revenue management department coordinates with the sales & marketing department and conducts social media marketing to the customers. By marketing to individual customers, such as sending sales campaign information and writing e-mail letters of thanks based on their attributes and their feedback on their last stay, the hotels gain customer loyalty.

4.2.2. Implications for stakeholders of hotels

Finally, our study has implications not only for hotels themselves but also for other stakeholders, namely shareholders, customers, suppliers, and media. Since supports from the relevant stakeholders on implementation of the revenue management model enhance revenues of the hotels and create a positive beneficial cycle for both parties, we provide some recommendations on how they can back up the hotels' revenue management.

- Shareholders: Shareholders should understand the hotel's efforts in revenue management by introducing the up-to-date revenue management model, which contribute to increase the hotel's firm value and the stock price. Then, the shareholders support the hotel's businesses by holding stocks, which also leads to the increase in the stock price and benefits both the shareholders and the hotel.

- Customers: The customers should pay close attention to the prices, reviews, and ratings of hotels in the online booking systems when they book a room. Additionally, they become keen to hotels' social media marketing, and build loyalty to the hotels that make efforts in revenue management and social media marketing. As a result, the revenues of such hotels grow, and in turn, these hotels are then able to provide the customers better services, which leads to a mutually beneficial relationship.

- Suppliers: Suppliers, such as housekeeping agencies and restaurants, should obtain the information of the expected number of customers from the hotels. Then, they can reduce food waste or surplus workforces, which increases the revenues of the suppliers, and they can then provide better services to the customers. As a result, the hotel gets more positive reviews from the customers, which creates a growth cycle.

  Online travel agencies should preferentially display the hotels, which make efforts in revenue management and social media marketing, on the top of their list in their website, and share helpful information on the customers to the hotels. Then, the booking probability of the hotels increases, which leads to revenue growth of the hotels and a fee increase of the online travel agencies.

- Media: Social media companies should ask hotels, which make efforts in revenue management and social media marketing, to place advertisements on their media. They display content depending on the attributes of the customers. Then, social media-induced bookings increase and the hotels have more incentive to place advertisements on social media. Similarly, internet television providers should also approach the hotels, which are keen on revenue management and social media marketing, and broadcast commercials and similar contents highlighting such features as visually impressive foods and facilities that can be spread by word of mouth or via social media. Then, the advertising effect increases, which leads to revenue growth of the hotels, and the hotels invest more into internet commercials and social media marketing.

5. Concluding Remarks

In this study, we have presented an application of online booking data, which are big data crawled from a hotel booking website, to hotel revenue management. Combining the choice behaviors of the customers estimated from the data by a discrete choice model and a quantitative overbooking model for online booking systems, we have investigated the optimal room charge and overbooking level that maximize the expected sales of the hotel. A revenue management for online booking systems that considers overbooking is particularly important since increasing last-minute cancellations can cause serious damage to revenues.

Moreover, we have provided the numerical examples of the optimal room charges and overbooking levels by using actual online booking data crawled from a Japanese booking website for two major luxury hotels in Shinjuku ward in Tokyo. We have obtained concrete optimal room charges and overbooking levels along with the maximized expected sales in the four different cases of the cancellation rate and over-sale cost per room.

Specifically, the results indicate concrete levels for the room charge and the overbooking, which cannot be obtained from the qualitative observation. For instance, when the over-sale cost per room and the cancellation rate are both high, the result exhibits the explicit trade-off between the over-sale cost and the compensation for the sales loss by the overbooking. In such a case, the model largely helps determine the overbooking level and the room charge.

Furthermore, in the numerical example of the equilibrium room charges of the two hotels, we have investigated the transition of the room charges when the hotels iteratively optimize them in turns. We have observed that the prices converge to lower levels compared to the original room charges and the expected sales accordingly drop.

5.1. Limitations and future research directions

As we mentioned in Remark 1 in Section 2.1, one limitation of our study is that the intensity of the total demand is independent of the overall levels of the room charges of the hotels in the area. Extending the model to the case where the intensity depends on the room charges will be one of our research topics.

Another limitation of our study is that the static factors related to social media and social media marketing, which the online booking customers consider when they book a room, are reflected only as a constant term of the preferences of hotels, but not as individual terms in the mixed logit model. Incorporating these factors into our modeling will further strengthen our model.

Online booking customers take into consideration not only room charges but also ratings and online reviews of hotels on social media after guests have booked a room. Additionally, social media marketing, such as advertisement and sales information individually displayed on the social media, affects the choice behaviors of the customers. For instance, Kapoor et al. (2018), Shiau et al. (2017), Shiau et al. (2018), Alalwan et al. (2017), and Dwivedi et al. (2015), summarize the studies on social media, social networks, and social media marketing. Shiau et al. (2017) conduct co-citation analysis for articles on social networks published between 1996 and 2014. Shiau et al. (2018) identify six points of core knowledge, including the behavioral analysis of users and the social impact and strategies of social networks, by the citation analysis for the studies on Facebook. Alalwan et al. (2017) review studies on the social media marketing, whose topics include the roles of social media on advertising, electronic word of mouth, and customer relationship management. Dwivedi et al. (2015) present a review of 71 articles on social media marketing, which connects the various aspects of the rapidly growing media marketing form. Kapoor et al. (2018) identify multiple emerging themes in the existing studies on social media and social networking from 1997 to 2017.

Although those factors are reflected in the constant term of the preference to the hotels in the mixed logit model in our model, it is not decomposed into individual factors. For example, it is possible to collect concrete information from online booking customers by conducting questionnaires on how social media or social media marketing affected them in their decision-making process. Together with the data on social media and social media marketing, it may be possible to construct a quantitative revenue management model that incorporates the effects of them on the choice behaviors of the customers. Namely, it would be possible to know how much expected sales increase by when considering the factors of social media and social media marketing, including online ratings and reviews as well as advertisements and social media (e.g., Twitter, Facebook, Instagram) reputations.

Consequently, advancement of studies on social media, social media marketing, and social media networks clarified by Kapoor et al. (2018), Shiau et al. (2017), Shiau et al. (2018), Alalwan et al. (2017), and Dwivedi et al. (2015) enables us to construct a revenue management model that incorporates social media and social media marketing factors into the model developed in this paper. This will be one of our main next research topics.

References

[1] Abrate, G., & Viglia, G. (2016). Strategic and tactical price decisions in hotel revenue management. *Tourism Management*, 55, 123-132.

[2] Alalwan, A. A., Rana, N. P., Dwivedi, Y. K., & Algharabat, R. (2017). Social media in marketing: A review and analysis of the existing literature. Telematics and Informatics34, 1177-1190.

[3] Ali, F., Kwak, K. S., & Kim, Y. G. (2016). Opinion mining based on fuzzy domain ontology and Support Vector Machine: A proposal to automate online review classification. *Applied Soft Computing*, 47, 235-250.

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

[5] Bilgihan, A., Nusair, K., Okumus, F., & Cobanoglu, C. (2015). Applying flow theory to booking experiences: An integrated model in an online service context. *Information & Management*, 52(6), 668-678.

[6] Chang, Y. C., Ku, C. H., & Chen, C. H. (2017). Social media analytics: Extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor. *International Journal of Information Management.*

[7] Chen, C. C., Schwartz, Z., & Vargas, P. (2011). The search for the best deal: How hotel cancellation policies affect the search and booking decisions of deal-seeking customers. *International Journal of Hospitality Management*, 30(1), 129-135.

[8] Chen, C. C., & Xie, K. L. (2013). Differentiation of cancellation policies in the US hotel industry. *International Journal of Hospitality Management*, 34, 66-72.

[9] Cheng, H., Lu, Y. C., & Chung, J. T. (2010). Improved slack-based context-dependent DEA - A study of international tourist hotels in Taiwan. *Expert Systems with Applications*, 37(9), 6452-6458.

[10] DeKay, F., Yates, B., & Toh, R. S. (2004). Non-performance penalties in the hotel industry. *International Journal of Hospitality Management*, 23(3), 273-286.

[11] Delgado, P. (2016). Cancellations on Booking.com: 104% more than on the hotel website. Expedia, 31% more. *mirai.com website at https://www.mirai.com/blog/cancellations-on-booking-com-104-more-than-on-the-hotel-website-expedia-31-more/ (accessed 11.01.2018)*

[12] Duro, J. A. (2016). Seasonality of hotel demand in the main Spanish provinces: Measurements and decomposition exercises. *Tourism Management*, 52, 52-63.

[13] Dwivedi, Y. K., Kapoor, K. K., & Chen, H. (2015). Social media marketing and advertising. The Marketing Review, 15(3), 289-309.

[14] Ferguson, M., & Smith, S. (2014). The changing landscape of hotel revenue management and the role of the hotel revenue manager. *Journal of Revenue and Pricing Management*, 13(3), 224-232.

[15] Gallego, G., & Van Ryzin, G. (1994). Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management science*, 40(8), 999-1020.

[16] Garrido-Moreno, A., & Padilla-Melendez, A. (2011). Analyzing the impact of knowledge management on CRM success: The mediating effects of organizational factors. *International Journal of Information Management*, 31(5), 437-444.

[17] Gehrels, S., & Blanar, O. (2013). How economic crisis affects revenue management: the case of the Prague Hilton hotels. *Research in Hospitality Management*, 2(1-2), 9-15.

[18] Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism management*, 23(5), 499-510.

[19] El Haddad, R. (2015). Exploration of revenue management practices?case of an upscale budget hotel chain. *International Journal of Contemporary Hospitality Management*, 27(8), 1791-1813.

[20] Hernandez-Ortega, B. (2018). Don ’ t believe strangers: Online consumer reviews and the role of social psychological distance. *Information & Management*, 55(1), 31-50.

[21] Hu, Y. H., & Chen, K. (2016). Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel stars and review ratings. *International Journal of Information Management*, 36(6), 929-944.

[22] Hu, Y. H., Chen, K., & Lee, P. J. (2017). The effect of user-controllable filters on the prediction of online hotel reviews. *Information & Management*, 54(6), 728-744.

[23] Huang, Y., Li, C., Wu, J., & Lin, Z. (2018). Online customer reviews and consumer evaluation: The role of review font. *Information & Management*, 55(4), 430-440.

[24] Horiguchi, H. (2015). Online group reservations and increasing cancellations, how hotels should prevent and cope with the problem (in Japanese). *travelvoice.jp website at https://www.travelvoice.jp/20150428-41369 (accessed 11.01.2018)*

[25] Ivanov, S., & Ayas, C. (2017). Investigation of the revenue management practices of accommodation establishments in Turkey: An exploratory study. *Tourism management perspectives*, 22, 137-149.

[26] Kapoor, K. K., Tamilmani, K., Rana, N. P., Patil, P., Dwivedi, Y. K., & Nerur, S. (2018). Advances in social media research: past, present and future. Information Systems Frontiers, 20(3), 531-558.

[27] Kardaras, D. K., Karakostas, B., & Mamakou, X. J. (2013). Content presentation personalisation and media adaptation in tourism web sites using Fuzzy Delphi Method and Fuzzy Cognitive Maps. *Expert Systems with Applications*, 40(6), 2331-2342.

[28] Kim, J. Y., Hlee, S., & Joun, Y. (2016). Green practices of the hotel industry: Analysis through the windows of smart tourism system. *International Journal of Information Management*, 36(6), 1340-1349.

[29] Koide, T., & Ishii, H. (2005). The hotel yield management with two types of room prices, overbooking and cancellations. *International Journal of Production Economics*, 93, 417-428.

[30] Korfiatis, N., & Poulos, M. (2013). Using online consumer reviews as a source for demographic recommendations: A case study using online travel reviews. *Expert Systems with Applications*, 40(14), 5507-5515.

[31] Liberman, V., & Yechiali, U. (1978). On the hotel overbooking problem - An inventory system with stochastic cancellations. *Management Science*, 24(11), 1117-1126.

[32] Mandelbaum, R. (2016). How attrition, cancellation fees hit your bottom line. *hotelnewsnow.com website at http://www.hotelnewsnow.com/Articles/54143/How-attrition-cancellation-fees-hit-your-bottom-line (accessed 11.01.2018)*

[33] Neirotti, P., Raguseo, E., & Paolucci, E. (2016). Are customers’ reviews creating value in the hospitality industry? Exploring the moderating effects of market positioning. *International Journal of Information Management*, 36(6), 1133-1143.

[34] Ozturk, A. B., Bilgihan, A., Nusair, K., & Okumus, F. (2016). What keeps the mobile hotel booking users loyal? Investigating the roles of self-efficacy, compatibility, perceived ease of use, and perceived convenience. *International Journal of Information Management*, 36(6), 1350-1359.

[35] Raguseo, E., Neirotti, P., & Paolucci, E. (2017). How small hotels can drive value their way in infomediation. The case of ‘Italian hotels vs. OTAs and TripAdvisor’. *Information & Management*, 54(6), 745-756.

[36] Rothstein, M. (1974). Hotel overbooking as a Markovian sequential decision process. *Decision Sciences*, 5(3), 389-404.

[37] Sierag, D. D., Koole, G. M., van der Mei, R. D., van der Rest, J. I., & Zwart, B. (2015). Revenue management under customer choice behaviour with cancellations and overbooking. *European Journal of Operational Research*, 246(1), 170-185.

[38] Sierag, D., Rest, J. P. V. D., Koole, G., Mei, R. V. D., & Zwart, B. (2017). A call for exploratory data analysis in revenue management forecasting: a case study of a small and independent hotel in The Netherlands. *International Journal of Revenue Management*, 10(1), 28-51.

[39] Shiau, W. L., Dwivedi, Y. K., & Yang, H. S. (2017). Co-citation and cluster analyses of extant literature on social networks. International Journal of Information Management, 37(5), 390-399.

[40] Shiau, W. L., Dwivedi, Y. K., & Lai, H. H. (2018). Examining the core knowledge on facebook. International Journal of Information Management, 43, 52-63.

[41] Shuai, J. J., & Wu, W. W. (2011). Evaluating the influence of E-marketing on hotel performance by DEA and grey entropy. *Expert Systems with Applications*, 38(7), 8763-8769.

[42] Takizawa, N. (2017). Realities of hotel reservation, a big problem of cancellations at the last minute (in Japanese). *bunshun.jp website at http://bunshun.jp/articles/-/2370 (accessed 11.01.2018)*

[43] Toh, R. S. (1986). Coping with no-shows, late cancellations and over-sales: American hotels out-do the airlines. *International Journal of Hospitality Management*, 5(3), 121-125.

[44] Toh, R. S., & Dekay, F. (2002). Hotel room-inventory management: an overbooking model. *The Cornell Hotel and Restaurant Administration Quarterly*, 43(4), 79-90.

[45] Train, K. E. (2009). Discrete choice methods with simulation. *Cambridge university press.*

[46] Tseng, M. L. (2009). Using the extension of DEMATEL to integrate hotel service quality perceptions into a cause?effect model in uncertainty. *Expert systems with applications*, 36(5), 9015-9023.

[47] Viglia, G., Minazzi, R., & Buhalis, D. (2016). The influence of e-word-of-mouth on hotel occupancy rate. *International Journal of Contemporary Hospitality Management*, 28(9), 2035-2051.

[48] Walls, M. (2016). 6 tips for combating high cancellation rates on OTA bookings. *netaffinity.com website at https://blog.netaffinity.com/6-tips-high-cancellation-rates-ota-bookings/ (accessed 11.01.2018).*

[49] Yu, S. M., Wang, J., Wang, J. Q., & Li, L. (2017). A multi-criteria decision-making model for hotel selection with linguistic distribution assessments. *Applied Soft Computing*, 67, 741-755.

[50] Xu, X., Wang, X., Li, Y., & Haghighi, M. (2017). Business intelligence in online customer textual reviews: Understanding consumer perceptions and influential factors. *International Journal of Information Management*, 37(6), 673-683.

Appendix A. Proof of Theorem 1

Let

$$f(x^{(i)}) = \mathbf{E}\left[\min\left(L_i^{act}, (1-r_i)R_T^{(i)}\right) x^{(i)}\right] - \mathbf{E}\left[\max\left((1-r_i)R_T^{(i)} - L_i^{act}, 0\right) c_i\right]. \tag{A.1}$$

First, we show that for a fixed $L_i^{ob} \geq L_i^{act}$, the maximum point with respect to $x^{(i)} \in (0, \infty)$ exists. To show this, it suffices to prove the following.

1.
$$\lim_{x^{(i)} \to 0} f(x^{(i)}) \leq 0, \tag{A.2}$$

2.
$$\lim_{x^{(i)} \to +\infty} f(x^{(i)}) = 0, \tag{A.3}$$

3. There exists $\bar{x}^{(i)} \in (0, \infty)$ such that
$$f(\bar{x}^{(i)}) > 0, \tag{A.4}$$

4. $f(x^{(i)})$ is a continuous function with respect to $x^{(i)}$.

Let $\gamma^* = (1, 1, \ldots, 1) \in \Gamma$ and $\gamma^\# \in \Gamma$ be

$$\gamma_j^\# = \begin{cases} 1, & j = i, J \\ 0, & \text{otherwise}, \end{cases} \qquad j = 1, \cdots, J.$$

1. $\lim_{x^{(i)} \to 0} f(x^{(i)}) \leq 0$

Noting that

$\lim_{x^{(i)} \to 0} f(x^{(i)})$

$= -c_i \lim_{x^{(i)} \to 0} \mathbf{E}\left[\max\left((1-r_i)R_T^{(i)} - L_i^{act}, 0\right)\right]$, we have $\lim_{x^{(i)} \to 0} f(x^{(i)}) \leq 0$,

since $-c_i \mathbf{E}\left[\max\left((1-r_i)R_T^{(i)} - L_i^{act}, 0\right)\right]$ is decreasing as $x^{(i)}$ decreases to 0 and satisfies

$$0 \geq -c_i \mathbf{E}\left[\max\left((1-r_i)R_T^{(i)} - L_i^{act}, 0\right)\right] \geq -c_i \mathbf{E}\left[\max\left((1-r_i)\min(N_T, L_i^{ob}) - L_i^{act}, 0\right)\right]. \tag{A.5}$$

2. $\lim_{x^{(i)} \to +\infty} f(x^{(i)}) = 0$

For the second term of $f(x^{(i)})$ in (A.1), since

$$0 \leq \max\left((1-r_i)R_T^{(i)} - L_i, 0\right) \leq R_T^{(i)}, \tag{A.6}$$

we have

$$\begin{aligned}
0 &\leq \mathbf{E}\left[\max((1-r_i)R_T^{(i)} - L_i^{act}, 0)\right] \\
&\leq \mathbf{E}[R_T^{(i)}] = \sum_{k=0}^{\infty} \mathbf{E}[R_T^{(i)} | N_T = k] \frac{(\lambda T)^k}{k!} e^{-\lambda T} \\
&\leq \sum_{k=0}^{\infty} \left(\sum_{m=0}^{k} m \binom{k}{m} p_i^{(\gamma^\#)m} (1 - p_i^{(\gamma^\#)})^{k-m} \right) \frac{(\lambda T)^k}{k!} e^{-\lambda T} \\
&= e^{-\lambda T} \sum_{m=0}^{\infty} m p_i^{(\gamma^\#)m} \left(\sum_{k=m}^{\infty} \binom{k}{m} (1 - p_i^{(\gamma^\#)})^{k-m} \frac{(\lambda T)^k}{k!} \right) \\
&= e^{-\lambda T} \sum_{m=0}^{\infty} m p_i^{(\gamma^\#)m} \left(\sum_{k'=0}^{\infty} \binom{k'+m}{m} (1 - p_i^{(\gamma^\#)})^{k'} \frac{(\lambda T)^{k'+m}}{(k'+m)!} \right) \\
&= e^{-\lambda T} \sum_{m=0}^{\infty} m p_i^{(\gamma^\#)m} \frac{(\lambda T)^m}{m!} \left(\sum_{k'=0}^{\infty} \frac{(\lambda T(1 - p_i^{(\gamma^\#)}))^{k'}}{k'!} \right) \\
&= e^{-\lambda T} e^{\lambda T(1 - p_i^{(\gamma^\#)})} \sum_{m=0}^{\infty} m p_i^{(\gamma^\#)m} \frac{(\lambda T)^m}{m!} \\
&= \lambda T p_i^{(\gamma^\#)} e^{-\lambda T p_i^{(\gamma^\#)}} \sum_{m'=0}^{\infty} \frac{(\lambda T p_i^{(\gamma^\#)})^{m'}}{m'!} \\
&= \lambda T p_i^{(\gamma^\#)}.
\end{aligned} \tag{A.7}$$

Hence,

$$\lim_{x^{(i)} \to +\infty} \mathbf{E}\left[\max\left((1-r_i)R_T^{(i)} - L_i^{act}, 0\right)\right] = 0. \tag{A.8}$$

Similarly, for the first term of $f(x^{(i)})$ in (A.1), since

$$0 \leq \mathbf{E}\left[\min\left(L_i^{act}, (1-r_i)R_T^{(i)}\right) x^{(i)}\right] \leq \mathbf{E}\left[R_T^{(i)} x^{(i)}\right] \tag{A.9}$$

and

$$\lim_{x^{(i)}\to\infty} \mathbf{E}\left[R_T^{(i)} x^{(i)}\right] \le \lim_{x^{(i)}\to\infty} \lambda T x^{(i)} p_i^{(\gamma^{\#})} = 0, \tag{A.10}$$

where the last equation will be shown later in the case of the mixed logit model (6), we have

$$\lim_{x^{(i)}\to\infty} \mathbf{E}\left[\min\left(L_i^{act}, (1-r_i)R_T^{(i)}\right) x^{(i)}\right] = 0. \tag{A.11}$$

Hence

$$\lim_{x^{(i)}\to+\infty} f(x^{(i)}) = 0. \tag{A.12}$$

To show (A.10) in the case of the mixed logit model in (6), noting that

$$p_i^{(\gamma^{\#})} x^{(i)} = \int_0^\infty x^{(i)} \frac{e^{-\beta x^{(i)}+\alpha_i}}{\sum_{j\in\{i,J\}} e^{-\beta x^{(j)}+\alpha_j}} h(\beta) d\beta, \tag{A.13}$$

and for $\beta > 0$, setting

$$x^{(*)} = x^{(J)},\ C \equiv e^{(\alpha_J-\alpha_i)} > 0, \tag{A.14}$$

we have

$$\begin{aligned}
\frac{x^{(i)}}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}} &= \frac{x^{(i)}}{1+Ce^{\beta(x^{(i)}-x^{(*)})}} \\
&= \frac{x^{(i)}-x^{(*)}}{1+Ce^{\beta(x^{(i)}-x^{(*)})}} + \frac{x^{(*)}}{1+Ce^{\beta(x^{(i)}-x^{(*)})}} \\
&\le \frac{x^{(i)}-x^{(*)}}{Ce^{\beta(x^{(i)}-x^{(*)})}} + x^{(*)} \le \frac{1}{Ce\beta} + x^{(*)}.
\end{aligned} \tag{A.15}$$

With the assumption that

$$\int_0^\infty \frac{1}{\beta} h(\beta) d\beta < \infty, \tag{A.16}$$

we have

$$\lim_{x^{(i)}\to\infty} p_i^{(\gamma^{\#})} x^{(i)} = 0 \tag{A.17}$$

by the dominated convergence theorem.

3. The existence of $\bar{x}^{(i)} \in (0,\infty)$ such that $f(\bar{x}^{(i)}) > 0$

To prove (A.4), it suffices to show that

$$\lim_{x^{(i)}\to\infty} \frac{\mathbf{E}\left[\max\left((1-r_i)R_T^{(i)} - L_i^{act}, 0\right) c_i\right]}{\mathbf{E}\left[\min\left(L_i^{act}, (1-r_i)R_T^{(i)}\right) x^{(i)}\right]} = 0.$$

This follows from the fact that

$$\begin{aligned}
\frac{\mathbf{E}\left[\max\left((1-r_i)R_T^{(i)} - L_i^{act}, 0\right) c_i\right]}{\mathbf{E}\left[\min\left(L_i^{act}, (1-r_i)R_T^{(i)}\right) x^{(i)}\right]} &\le \frac{c_i \mathbf{E}\left[R_T^{(i)}\right]}{x^{(i)}(1-\bar{r}_i)\mathbf{P}\left(R_T^{(i)}=1\right)} \\
&\le \frac{c_i \lambda T p_i^{(\gamma^{\#})}}{x^{(i)}(1-\bar{r}_i)\frac{\lambda T p_i^{(\gamma^*)}}{e^{\lambda T p_i^{(\gamma^*)}}}} \le \frac{c_i e^{\lambda T} p_i^{(\gamma^{\#})}}{(1-\bar{r}_i)x^{(i)} p_i^{(\gamma^*)}} \to 0\ (x^{(i)}\to\infty).
\end{aligned} \tag{A.18}$$

Here, we used

$$
\begin{aligned}
\mathbf{P}\left(R_T^{(i)}=1\right) &= \sum_{k=0}^{\infty}\mathbf{P}\left(R_T^{(i)}=1|N_T=k\right)\mathbf{P}(N_T=k) \\
&\geq \sum_{k=0}^{\infty} k p_i^{(\gamma^*)}\left(1-p_i^{(\gamma^*)}\right)^{k-1}\frac{e^{-\lambda T}}{k!}(\lambda T)^k \\
&= \lambda T p_i^{(\gamma^*)} e^{-\lambda T}\sum_{k'=0}^{\infty}\left(1-p_i^{(\gamma^*)}\right)^{k'}\frac{(\lambda T)^{k'}}{k'!} \\
&= \lambda T p_i^{(\gamma^*)} e^{-\lambda T} e^{\lambda T\left(1-p_i^{(\gamma^*)}\right)} = \lambda T p_i^{(\gamma^*)} e^{-\lambda T p_i^{(\gamma^*)}},
\end{aligned} \tag{A.19}
$$

and

$$
\lim_{x^{(i)}\to\infty}\frac{p_i^{(\gamma^\#)}}{x^{(i)}p_i^{(\gamma^*)}} = 0, \tag{A.20}
$$

which is shown particularly in the case of the mixed logit model in (6) as follows.

For $\bar{\beta}>0$,

$$
\begin{aligned}
\frac{p_i^{(\gamma^\#)}}{x^{(i)}p_i^{(\gamma^*)}} &= \frac{\int_0^\infty \frac{1}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}h(\beta)d\beta}{x^{(i)}\int_0^\infty \frac{1}{1+\sum_{j\neq i}e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}}h(\beta)d\beta} \\
&\leq \frac{\int_0^{\bar{\beta}} \frac{1}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}h(\beta)d\beta}{x^{(i)}\int_0^{\bar{\beta}} \frac{1}{1+\sum_{j\neq i}e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}}h(\beta)d\beta} + \frac{\int_{\bar{\beta}}^\infty \frac{1}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}h(\beta)d\beta}{x^{(i)}\int_0^{\bar{\beta}} \frac{1}{1+\sum_{j\neq i}e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}}h(\beta)d\beta}.
\end{aligned} \tag{A.21}
$$

For the first term on the right-hand side of (A.21), for $0\leq\beta\leq\bar{\beta}$,

$$
\begin{aligned}
\frac{1}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}} &= \frac{1}{1+\sum_{j\neq i}e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}}\frac{1+\sum_{j\neq i}e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}} \\
&= \frac{1}{1+\sum_{j\neq i}e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}}\left(1+\frac{\sum_{j\neq i,J}e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}\right) \\
&= \frac{1}{1+\sum_{j\neq i}e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}}\left(1+\frac{e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}\sum_{j\neq i,J}e^{-\beta(x^{(j)}-x^{(J)})+(\alpha_j-\alpha_J)}\right) \\
&\leq \frac{1}{1+\sum_{j\neq i}e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}}\left(1+\left(\sum_{j\neq i,J}e^{(\alpha_j-\alpha_J)}\right)\max_{j\neq i,J,0\leq\beta\leq\bar{\beta}}e^{-\beta(x^{(j)}-x^{(J)})}\right).
\end{aligned} \tag{A.22}
$$

Hence

$$
\begin{aligned}
&\int_0^{\bar{\beta}}\frac{1}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}h(\beta)d\beta \\
&\leq \left(\int_0^{\bar{\beta}}\frac{1}{1+\sum_{j\neq i}e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}}h(\beta)d\beta\right)\left(1+\left(\sum_{j\neq i,J}e^{(\alpha_j-\alpha_J)}\right)\max_{j\neq i,J,0\leq\beta\leq\bar{\beta}}e^{-\beta(x^{(j)}-x^{(J)})}\right)
\end{aligned} \tag{A.23}
$$

and

$$\frac{\int_0^{\bar\beta} \frac{1}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}} h(\beta)d\beta}{x^{(i)}\int_0^{\bar\beta} \frac{1}{1+\sum_{j\neq i} e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}} h(\beta)d\beta} \leq \frac{1}{x^{(i)}}\left(1+\left(\sum_{j\neq i,J} e^{(\alpha_j-\alpha_J)}\right)\max_{j\neq i,J,0\leq\beta\leq\bar\beta} e^{-\beta(x^{(j)}-x^{(J)})}\right)$$
$$\to 0 \ (x^{(i)}\to\infty). \tag{A.24}$$

For the second term in the right hand side of (A.21), for $\beta \geq \bar\beta$ and $x^{(i)} > \max_{j\neq i} x^{(j)}$, since

$$\frac{1}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}} \leq \frac{1}{1+e^{-\bar\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}, \tag{A.25}$$

we have

$$\int_{\bar\beta}^{\infty} \frac{1}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}} h(\beta)d\beta \leq \frac{1}{1+e^{-\bar\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}. \tag{A.26}$$

Hence,

$$\begin{aligned}
\frac{\int_{\bar\beta}^{\infty} \frac{1}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}} h(\beta)d\beta}{x^{(i)}\int_0^{\bar\beta} \frac{1}{1+\sum_{j\neq i} e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}} h(\beta)d\beta} &\leq \frac{1}{x^{(i)}\int_0^{\bar\beta} \frac{1+e^{-\bar\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}{1+\sum_{j\neq i} e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}} h(\beta)d\beta} \\
&\leq \frac{1}{x^{(i)}\int_0^{\bar\beta} \frac{1+e^{-\bar\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}{1+\sum_{j\neq i} e^{-\bar\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}} h(\beta)d\beta} \\
&= \frac{1}{x^{(i)}\left(\int_0^{\bar\beta} h(\beta)d\beta\right)\frac{1+e^{-\bar\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}{1+\sum_{j\neq i} e^{-\bar\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}}}.
\end{aligned} \tag{A.27}$$

Since

$$\lim_{x^{(i)}\to\infty} \frac{1+e^{-\bar\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}{1+\sum_{j\neq i} e^{-\bar\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}} = \frac{e^{-\bar\beta x^{(J)}+(\alpha_J-\alpha_i)}}{\sum_{j\neq i} e^{-\bar\beta x^{(j)}+(\alpha_j-\alpha_i)}}, \tag{A.28}$$

we have

$$\frac{\int_{\bar\beta}^{\infty} \frac{1}{1+e^{-\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}} h(\beta)d\beta}{x^{(i)}\int_0^{\bar\beta} \frac{1}{1+\sum_{j\neq i} e^{-\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}} h(\beta)d\beta} \leq \frac{1}{x^{(i)}\left(\int_0^{\bar\beta} h(\beta)d\beta\right)\frac{1+e^{-\bar\beta(x^{(J)}-x^{(i)})+(\alpha_J-\alpha_i)}}{1+\sum_{j\neq i} e^{-\bar\beta(x^{(j)}-x^{(i)})+(\alpha_j-\alpha_i)}}}$$
$$\to 0 \ (x^{(i)}\to\infty). \tag{A.29}$$

Therefore,

$$\lim_{x^{(i)}\to\infty} \frac{p_i^{(\gamma^{\#})}}{x^{(i)} p_i^{(\gamma^*)}} = 0. \tag{A.30}$$

4. The continuity of $f(x^{(i)})$

We first show the continuity of the first term of $f(x^{(i)})$ in (A.1).
Set

$$g(x^{(i)}) = \mathbf{E}\left[\min(L_i^{act}, (1-r_i)R_T^{(i)})\right], \tag{A.31}$$

$$g_N(x^{(i)}) = \sum_{k=0}^{N} \min(L_i^{act}, (1-r_i)\min(k, L_i^{ob}))\mathbf{P}(R_T^{(i)} = k). \tag{A.32}$$

Then,

$$\lim_{N\to\infty} \sup_{x^{(i)}\in(0,\infty)} |g_N(x^{(i)}) - g(x^{(i)})| = 0, \tag{A.33}$$

since

$$\begin{aligned}&\sup_{x^{(i)}\in(0,\infty)} |g_N(x^{(i)}) - g(x^{(i)})| \le \sup_{x^{(i)}\in(0,\infty)} \sum_{k=N+1}^{\infty} k\mathbf{P}(R_T^{(i)} = k)\\ &= \sup_{x^{(i)}\in(0,\infty)} \mathbf{E}\left[R_T^{(i)} 1_{\{R_T^{(i)}\ge N+1\}}\right] \le \mathbf{E}\left[N_T 1_{\{N_T\ge N+1\}}\right] \to 0,\ (N\to\infty).\end{aligned} \tag{A.34}$$

Hence, $g(x^{(i)})$ is a continuous function, since the continuous function $g_N(x^{(i)})$ converges to $g(x^{(i)})$ uniformly in $x^{(i)} \in (0,\infty)$, and the continuity of $x^{(i)}g(x^{(i)})$ follows.

In the same manner, it follows that the second term of $f(x^{(i)})$ in (A.1),

$$\mathbf{E}\left[-\max((1-r_i)R_T^{(i)} - L_i^{act}, 0)c_i\right] \tag{A.35}$$

is a continuous function with respect to $x^{(i)}$. Hence, $f(x^{(i)})$ is continuous on $(0,\infty)$.

Finally, for fixed $x^{(i)} \in (0,\infty)$,

$$\mathbf{E}\left[\min\left(L_i^{act}, (1-r_i)R_T^{(i)}\right)x^{(i)}\right] - \mathbf{E}\left[\max\left((1-r_i)R_T^{(i)} - L_i^{act}, 0\right)c_i\right];\ r_i \in [0,\bar{r}_i] \subset [0,1) \tag{A.36}$$

is strictly decreasing for $L_i^{ob} \ge \frac{L_i^{act}}{1-\bar{r}_i}$.

In fact, for $L_i^{ob} \ge \frac{L_i^{act}}{1-\bar{r}_i}$, if L_i^{ob} increases from $\bar{L}_i^{ob}$ to $\bar{L}_i^{ob}+1$, only the probabilities of $R_T^{(i)} = \bar{L}_i^{ob}$ and $R_T^{(i)} = \bar{L}_i^{ob}+1$ change. That is, for selection orders τ that contain $\bar{L}_i^{ob}+1$ of i, the value of $R_T^{(i)}$ changes from $\bar{L}_i^{ob}$ to $\bar{L}_i^{ob}+1$, and the probability for such selection orders changes from 0 to some positive value. For the first term $\mathbf{E}\left[\min\left(L_i^{act}, (1-r_i)R_T^{(i)}\right)x^{(i)}\right]$, since $\min\left(L_i^{act}, (1-r_i)R_T^{(i)}\right)x^{(i)} = L_i^{act}x^{(i)}$, it is unchanged. For the second term $\mathbf{E}\left[-\max\left((1-r_i)R_T^{(i)} - L_i^{act}, 0\right)c_i\right]$, since only for such selection orders τ, $-\max\left((1-r_i)R_T^{(i)} - L_i^{act}, 0\right)c_i$ decreases from $-\left((1-r_i)\bar{L}_i^{ob} - L_i^{act}\right)c_i$ to $-\left((1-r_i)(\bar{L}_i^{ob}+1) - L_i^{act}\right)c_i$, and the probability changes from 0 to some positive value, it is strictly decreasing, and hence (A.36) is strictly decreasing for $L_i^{ob} \ge \frac{L_i^{act}}{1-\bar{r}_i}$. Hence, (A.36) has a maximum at some $L_i^{ob} \in [L_i^{act},\infty)$, namely, $L_i^{ob} \in [L_i^{act}, L_i^{act}/(1-\bar{r}_i)]$.

Therefore,

$$\max_{L_i^{ob}\ge L_i^{act}} \max_{x^{(i)}\in(0,\infty)} \mathbf{E}[\min\left(L_i^{act}, (1-r_i)R_T^{(i)}\right)x^{(i)} - \max\left((1-r_i)R_T^{(i)} - L_i^{act}, 0\right)c_i]. \tag{A.37}$$

is attained at some $(x_*^{(i)}, L_{i*}^{ob})$ in $(0,\infty)\times[L_i^{act},\infty)$, which is also a maximum point for the objective function in (16). □

Appendix B. Proof of Theorem 2

Let

$$f(x_1) = x_1 p_1 = \frac{x_1}{1+e^{\beta(x_1-x_2)+\alpha_2-\alpha_1}}, \tag{B.1}$$

$$g(x_2) = x_2 p_2 = \frac{x_2}{1+e^{\beta(x_2-x_1)+\alpha_1-\alpha_2}}. \tag{B.2}$$

Since $\lim_{x_1 \to 0} f(x_1) = 0, \lim_{x_1 \to \infty} f(x_1) = 0$ and $f(x_1) \geq 0$ on $(0, \infty)$, $f(x_1)$ attains its maximum at some x_1 that satisfies $f'(x_1) = 0$.

Noting that

$$f'(x_1) = \frac{1 + e^{\beta(x_1 - x_2) + \alpha_2 - \alpha_1} - \beta x_1 e^{\beta(x_1 - x_2) + \alpha_2 - \alpha_1}}{\left(1 + e^{\beta(x_1 - x_2) + \alpha_2 - \alpha_1}\right)^2}, \tag{B.3}$$

we have

$$e^{\beta x_2} = (\beta x_1 - 1) e^{\beta x_1 + \alpha_2 - \alpha_1}, \tag{B.4}$$

and

$$x_2 = \frac{1}{\beta} \log(\beta x_1 - 1) + x_1 + \frac{\alpha_2 - \alpha_1}{\beta} = h(x_1). \tag{B.5}$$

Since $h : (\frac{1}{\beta}, \infty) \to (-\infty, \infty)$ is strictly increasing and hence bijective, there exists $h^{-1} : (-\infty, \infty) \to (\frac{1}{\beta}, \infty)$ such that

$$x_1 = h^{-1}(x_2). \tag{B.6}$$

Similarly, $g(x_2)$ attains its maximum at x_2 satisfying

$$e^{\beta x_1} = (\beta x_2 - 1) e^{\beta x_2 + \alpha_1 - \alpha_2} \tag{B.7}$$

or equivalently

$$x_1 = \frac{1}{\beta} \log(\beta x_2 - 1) + x_2 + \frac{\alpha_1 - \alpha_2}{\beta} = i(x_2), \tag{B.8}$$

and $i : (\frac{1}{\beta}, \infty) \to (-\infty, \infty)$ has the inverse $i^{-1} : (-\infty, \infty) \to (\frac{1}{\beta}, \infty)$ such that

$$x_2 = i^{-1}(x_1). \tag{B.9}$$

A Nash equilibrium is a solution of the simultaneous equations,

$$e^{\beta x_2} = (\beta x_1 - 1) e^{\beta x_1 + \alpha_2 - \alpha_1} \tag{B.10}$$

$$e^{\beta x_1} = (\beta x_2 - 1) e^{\beta x_2 + \alpha_1 - \alpha_2}. \tag{B.11}$$

Substituting (B.10) for (B.11), we have

$$1 = (\beta x_2 - 1)(\beta x_1 - 1), \tag{B.12}$$

and

$$1 = \left(\log(\beta x_1 - 1) + \beta x_1 + \alpha_2 - \alpha_1 - 1\right)(\beta x_1 - 1). \tag{B.13}$$

Setting $\beta x_1 - 1 = v$ and $\beta x_2 - 1 = w$, we have $v, w > 0$,

$$w = \frac{1}{v}, \tag{B.14}$$

and

$$1 = (\log v + v + \alpha_2 - \alpha_1) v. \tag{B.15}$$

Since the right hand side of (B.15) is strictly increasing, the simultaneous equations (B.10) and (B.11) have a unique solution $(x_*^{(1)}, x_*^{(2)}) \in (\frac{1}{\beta}, \infty)^2$.

Finally, we show that for any $(x_0^{(1)}, x_0^{(2)}) \in [0,\infty)^2$ and $F : [0,\infty)^2 \to [0,\infty)^2$ such that $F(x_1,x_2) = (h^{-1}(x_2), i^{-1}(x_1))$, $\{(x_n^{(1)}, x_n^{(2)})\}_{n\in\mathbf{N}}$ defined by
$(x_{n+1}^{(1)}, x_{n+1}^{(2)}) = F(x_n^{(1)}, x_n^{(2)})$ converges to $(x_*^{(1)}, x_*^{(2)})$ as $n \to \infty$.

Since for $h : (\frac{1}{\beta}, \infty) \to (-\infty,\infty)$,

$$\frac{dh(x_1)}{dx_1} = \frac{\beta}{\beta(\beta x_1 - 1)} + 1 = \frac{\beta x_1}{\beta x_1 - 1} \neq 0, \tag{B.16}$$

we have for $h^{-1} : (-\infty,\infty) \to (\frac{1}{\beta}, \infty)$,

$$\frac{dh^{-1}(x_2)}{dx_2} = \frac{\beta x_1 - 1}{\beta x_1} = \frac{\beta h^{-1}(x_2) - 1}{\beta h^{-1}(x_2)}, \tag{B.17}$$

and

$$0 < \frac{dh^{-1}(x_2)}{dx_2} < 1. \tag{B.18}$$

Consider a ball $B_R = \{(x_1,x_2) | |(x_1,x_2) - (x_*^{(1)}, x_*^{(2)})| \leq R\}$ with the radius $R > 0$ and the center $(x_*^{(1)}, x_*^{(2)})$. Here, we take $R > 0$ so that $(x_0^{(1)}, x_0^{(2)}) \in B_R$. We show that F restricted on B_R is a contraction map from $B_R \cap [0,\infty)^2$ to $B_R \cap [0,\infty)^2$.

For any $(x_1,x_2) \in B_R \cap [0,\infty)^2$, $F(x_1,x_2) = (h^{-1}(x_2), i^{-1}(x_1))$ and

$$\begin{aligned} |h^{-1}(x_2) - x_*^{(1)}| &= |h^{-1}(x_2) - h^{-1}(x_*^{(2)})| \\ &\leq \left| \int_{x_*^{(2)}}^{x_2} \frac{dh^{-1}(x_2)}{dx_2} dx_2 \right| \leq |x_2 - x_*^{(2)}|. \end{aligned} \tag{B.19}$$

Similarly,

$$\begin{aligned} |i^{-1}(x_1) - x_*^{(2)}| &= |i^{-1}(x_1) - i^{-1}(x_*^{(1)})| \\ &\leq \left| \int_{x_*^{(1)}}^{x_1} \frac{di^{-1}(x_1)}{dx_1} dx_1 \right| \leq |x_1 - x_*^{(1)}|. \end{aligned} \tag{B.20}$$

Hence, $F(x_1,x_2) \in B_R \cap [0,\infty)^2$.

Next, noting that $|(x_1,x_2)| \leq R + |(x_*^{(1)}, x_*^{(2)})|$ for all $(x_1,x_2) \in B_R \cap [0,\infty)^2$, from (B.17), we have

$$0 < \frac{dh^{-1}(x_2)}{dx_2} < 1 - \frac{1}{\beta(R + |(x_*^{(1)}, x_*^{(2)})|)}, \tag{B.21}$$

and similarly

$$0 < \frac{di^{-1}(x_1)}{dx_1} < 1 - \frac{1}{\beta(R + |(x_*^{(1)}, x_*^{(2)})|)}. \tag{B.22}$$

Hence, taking $0 < r_R = 1 - \frac{1}{\beta(R + |(x_*^{(1)}, x_*^{(2)})|)} < 1$, we have $|F(x_1,x_2) - F(\bar{x}_1,\bar{x}_2)| \leq r_R |(x_1,x_2) - (\bar{x}_1,\bar{x}_2)|, \forall (x_1,x_2), (\bar{x}_1,\bar{x}_2) \in B_R \cap [0,\infty)^2$, which implies that F is a contraction map from $B_R \cap [0,\infty)^2$ to $B_R \cap [0,\infty)^2$.

For $(x_0^{(1)}, x_0^{(2)}) \in B_R \cap [0,\infty)^2$, define $\{(x_n^{(1)}, x_n^{(2)})\}_{n\in\mathbf{N}}$ by

$$(x_{n+1}^{(1)}, x_{n+1}^{(2)}) = F(x_n^{(1)}, x_n^{(2)}). \tag{B.23}$$

Since F is a contraction map, there exists $(x_\infty^{(1)}, x_\infty^{(2)}) \in B_R \cap [0,\infty)^2$ such that

$$\lim_{n\to\infty} (x_n^{(1)}, x_n^{(2)}) = (x_\infty^{(1)}, x_\infty^{(2)}). \tag{B.24}$$

Noting that F is a continuous map, taking the limit $n \to \infty$ on

$$(x_{n+1}^{(1)}, x_{n+1}^{(2)}) = F(x_n^{(1)}, x_n^{(2)}), \tag{B.25}$$

we have

$$(x_\infty^{(1)}, x_\infty^{(2)}) = F(x_\infty^{(1)}, x_\infty^{(2)}), \tag{B.26}$$

which implies that $(x_\infty^{(1)}, x_\infty^{(2)})$ is the Nash equilibrium, and therefore $(x_\infty^{(1)} x_\infty^{(2)}) = (x_*^{(1)}, x_*^{(2)})$. $\square$