Overbooking for Physical Examination considering Late Cancellation and Set-Resource Relationship

Te-Wei Ho[1]**, Ling-Chieh Kung**[2]*****, Hsin-Ya Huang**[2]**, Jui-Fen Lai**[3]**, and Han-Mo Chiu**[3,4]

[1] Department of Surgery, College of Medicine, National Taiwan University, Taipei, Taiwan.

[2] Department of Information Management, College of Management, National Taiwan University, Taipei, Taiwan.

[3] Health Management Center, National Taiwan University Hospital, Taipei, Taiwan.

[4] Department of Internal Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan.

* Corresponding author: lckung@ntu.edu.tw, ORCID: 0000-0002-9346-0975.

Abstract

Background: Late cancellations of physical examination has severe impact on the operations of a physical examination center since it is often too late to fill vacancy. A booking control policy that considers overbooking is then one natural solution. Unlike appointment scheduling problems for clinics and hospitals, in which treating a patient mostly requires only one type of resource, a physical examination set typically requires multiple types of resources. Traditional methods that do not consider set-resource relationship thus may be inapplicable.
Methods: We formulate a stochastic mathematical programming model that maximizes the expected net reward, which is the examination revenue minus overage cost. A complete search algorithm and a greedy search algorithm are designed to search for optimal booking limits for all examination sets. To estimate the late cancellation probability for each individual consumer, we apply logistic regression to identify significant factors affecting the probability. After clustering is used to estimate individual probabilities, Monte Carlo simulation is conducted to generate probability distributions for the number of consumers without late cancellations. A discrete-event simulation is performance to evaluate the effectiveness of our proposed solution.
Results: We collaborate with a leading physical examination center to collect real data to evaluate our proposed overbooking policies. We show that the proposed overbooking policy may significantly increase the expected net reward. Our simulation results also help us understand the impact of overbooking on the expected number of customers and expected overage. A sensitivity analysis is conducted to demonstrate that the benefit of overbooking is insensitive to the accuracy of cost estimation. A Pareto efficiency analysis gives practitioners suggestions regarding policy determination considering multiple performance indications.
Conclusions: Our proposed overbooking policies may greatly enhance the overall performance of a physical examination center.

1 Background

1.1 Overbooking considering Sets and Resources

No-shows and late cancellations, which cause healthcare resource waste and longer patient/consumer waiting, are critical issues for healthcare service providers to deal with [5,26]. It is documented that the no-show rates in physical examination related sectors may be up to 17.6% [28,3,25]. In a survey with respect to 200 continuity clinics, the mean no-show percentage is 30.9%, where the maximum one is 80% [32]. In fact, among clinic size, no-show probability, and variation of service times, which are the three major environmental factors affecting the performance of a healthcare appointment system, no-show probability is considered the most significant [12]. It is widely suggested that no-shows and late cancellations should be taken into account when designing an appointment management system [11].

For healthcare environments with no-shows and late cancellations, operations researchers have done a lot of works in appointment management to enhance efficiency [7]. To deal with the uncertainty brought by random no-shows and late cancellation, one way is to adopt dynamic appointment scheduling [9,23,35]. By doing this, a healthcare service provider allocates physicians' available time slots to patients periodically by considering random appointment requests, no-shows, and late cancellations. While patients' preference over physicians and/or time slots and delay-dependent no-show probability are the major issues under consideration, the capacity (e.g., number of time slots in each period) is considered fixed. Another way to tackle no-shows and late cancellations is to relax the fixed capacity constraint and adopt overbooking, which has been widely applied in the hotel and airlines industry decades ago [24,31]. The possibility and potential benefit of overbooking in the healthcare industry are also investigated in recent years [1,4,15,18,22,21,16,27,33]. Given no-show probabilities of patients and an ideal number of patients for each time slot, these studies aim to determine a booking limit (which is typically larger than the ideal number of patients) to find a balance between clinic revenue, physician overtime, patient waiting, and the general accessibility of medical services.

For the physical examination center we collaborate with, dynamic appointment scheduling is already applied, and overbooking is under consideration. Interestingly, while there are many past works that develop overbooking policies for outpatients requesting clinical services, none of the policies developed in the above works may be directly applied to the physical examination center due to a major difference between these

two types of services. For most clinical services, each outpatient typically requires only one type of *resource* (e.g., a physician). When developing an overbooking policy, booking limits are determined directly for resources (e.g., at most five bookings for a specific physician in an hour), and whether *overage* occurs (i.e., more patients than expected show up) is determined also for each resource (e.g., whether a physician needs to work overtime). In short, the determination of booking limits and calculation of overage costs are aligned.

Physical examinations, however, are mostly provided to consumers in the format of *sets*, where an examination set requires multiple resources such as gastroenteroscope, magnetic resonance imaging (MRI), computed tomography scan (CT), etc. For example, our collaborating center offers 22 examination sets for consumers to select from (cf. Table 13 in the appendix for details). Different examination sets require different resources; for example, among the 22 examinations sets, three of them require gastroenteroscope, six require MRI, and two require CT. For most physical examination centers, while overage levels are still calculated based on resources, booking limits are determined for examination sets. In other words, the determination of booking limits and calculation of overage costs are *not* aligned. The fact that the same level of overage at different resources may bring in quite different costs and inefficiency further complicates the problem.

In summary, the relationship between sets and resources, which is called *set-resource relationship* in this study, must be incorporated in developing an overbooking policy for an physical examination center. To the best of our knowledge, this work is the first one that develops a mathematical model of overbooking considering the set-resource relationship. This work may thus contribute to both practice and academia.

1.2 Factors affecting Late Cancellation Probabilities

To generate an overbooking policy in practice, a theoretical model is not enough. The estimation of late cancellation probability based on historical booking and cancellation data is also required.

Many works have been devoted to the investigation of factors explaining patient no-shows, mostly about clinical appointments. Parente et al. [28] classify the explanatory variables of the early works into three groups: (1) factors related to individuals, (2) factors related to environmental conditions, and (3) factors associated with practices designed to increase patient's attendance rates. Regarding demographic information of patients, it has been reported that age, gender, ethnicity, type of medical insurance, socioeconomic condi-

tions, diseases and conditions, past no-show and late cancellation records are correlated with no-shows and late cancellations [6, 10, 14, 33]. For environmental factors, information about office accessibility, transportation convenience, whether childcare is provided, appointment lead time (appointment delay), examination type (specialty group), examination season, and service quality have been shown to be effective [2, 8, 10, 20, 28]. Finally, education, penalties, and reminders by phone or email are also documented as useful devices to reduce no-show probabilities. [13, 17, 19, 29, 32]. For most the works mentioned above, regression-type analysis is the major research tool.

In this study, we follow the above works to search for significant factors that are correlated with late cancellations in the collaborating physical examination center. While the above works provide some guidance, not all variables are available in this study. For example, the center has no information about a consumer's ethnicity and type of medical insurance. Moreover, as there is only one center in this study, there is no way to examine facility-related factors such as office accessibility and transportation convenience and activities like education and reminders using historical bookings. Table 1 lists the variables we obtain from the collaborating center that have been documented in the literature, including a consumer's age, a consumer's gender, appointment lead time (number of days between the examination day and appointment day), season (in which month the examination is conducted), and examination set. We investigate these variables about their significance and use the significant ones to construct our late cancellation probability estimation model.

There are also some variables that are not documented in literature but turn out to be effective in this study. In particular, we find that whether a booking is a group booking, i.e., more than one consumers book for the same day in one phone call or one online submission, and whether a booking is in a national holiday have significant impact on late cancellation probabilities. These variables are also considered to be used in the process of estimating late cancellation probabilities. In other words, this work adds to the literature by reporting new variables that are significantly correlated with late cancellations.

1.3 A Comprehensive Case Study

In this study, we construct a stochastic optimization model to do overbooking by considering the set-resource relationship. We also propose a way of late cancellation probability estimation to generate the estimated probability for each individual consumer. To further create

Table 1 Factors Comparison between This Study and the Literature

Variable	Our Study	[2]	[8]	[10]	[13]	[14]	[20]	[28]	[33]
Age	S.		v		v	v	v	v	v
Gender	N.S.							v	
Appointment lead time	S.	v	v	v			v		
Season	S.							v	
Examination set (specialty group)	N.S.	v						v	
Group booking	S.								
Holiday	S.								

S. means the variable is statistically significant and N.S. means the variable is not significant in our study.
v means the variable is statistically significant in that study.

values for physical examination practitioners, we collect real booking and cancellation data from our collaborating physical examination center to conduct a comprehensive case study. We describe the whole process of data cleansing, probability estimation, booking limit optimization, and performance evaluation of various overbooking policies. We believe that the description of the complete process may serve as a reference for practitioners to apply the same idea in their own physical examination centers.

For ease of exposition, in the sequel we will use the term "late cancellations" to include both late cancellations and no-shows.

2 Methods

To generate an overbooking policy, there are two major steps: late cancellation probability estimation and booking limit optimization, where the first part is to find parameter values for the second part. Below we will first introduce the stochastic mathematical programming model for booking limit optimization. We will then describe the process of late cancellation probability estimation, which is based on logistic regression and Monte Carlo simulation. While the general process is presented in this section, a detailed case study that may serve as an example may be found in the next section.

2.1 Stochastic Mathematical Programming

Consider a physical examination center which opens for booking and subject to late cancellations. The center provides various physical examination *sets* to fit different consumers' needs. Each set includes several *resources*, such as gastroenteroscope, MRI, CT, etc. Let $I = \{1, 2, ..., n\}$ be the collection of resources and $J = \{1, 2, ..., m\}$ be the collection of examination sets, where m is the number of examination sets. The original *booking limit* (numbers of open slots) and price of

set j are B_j and P_j, respectively, for each $j \in J$. Let X_j be the random number of consumers that show up in a day to do set j with no late cancellation and x_j be a possible realization of X_j. The *expected revenue* that the center may earn with the original booking limits may then be calculated as

$$\sum_{x_1=1}^{B_1} \sum_{x_2=1}^{B_2} \cdots \sum_{x_m=1}^{B_m} x_j P_j \prod_{j \in J} \Pr(X_j = x_j),$$

where $\Pr(X_j = x_j)$ is the probability for x_j consumers to show up for set j.

Knowing that the B_j consumers who booked set j may not all show up, the center manager is considering to increase the booking limit to N_j for set j. We will call a collection of booking limits $N_1, N_2, ...,$ and N_m as a *booking limit combination*, which is to be determined by the center manager. Obviously, in an overbooking policy we have $N_j \geq B_j$. The expected revenue that the center may earn with the booking limit combination $N_1, N_2, ...,$ and N_m, which is

$$\sum_{x_1=1}^{N_1} \sum_{x_2=1}^{N_2} \cdots \sum_{x_m=1}^{N_m} x_j P_j \prod_{j \in J} \Pr(X_j = x_j),$$

is obviously no less than that with the original booking limits as long as $N_j \geq B_j$ for all $j \in J$. However, each resource has its *ideal consumer level*, which is the maximum number of consumers that may be served with that resource in regular time. Once the number of arrived consumers requiring a resource exceed the ideal consumer level, for which we say *overage* occurs, overtime work is needed and consumer waiting cannot be avoided.

The *overage cost* under a booking limit combination $N_1, N_2, ...,$ and N_m may be expressed as follows. Let K_i be the ideal consumer level for resource i and V be an $n \times m$ matrix indicating the set-resource relationship, where $V_{ij} = 1$ if set j requires resource i and 0 otherwise. If x_j consumers who booked set j show up

in a day, The *overage level* of resource i is

$$\max\left\{\sum_{j\in J} V_{ij}x_j - K_i, 0\right\}.$$

When the overage level of a resource is positive, overtime work and consumer waiting incur some kind of cost for the physical examination center. Let $f_i(z)$ be the overage cost of resource i incurred by the center when the overage level is z. To reflect the practice of a physical examination center, $f_i(z)$ should be strictly increasing and strictly convex (so that the marginal overage cost is increasing). We should also have $f(z) = 0$ for all $z \leq 0$. The total overage cost across all resources is thus

$$\sum_{i\in I} f_i\left(\sum_{j\in J} V_{ij}x_j - K_i\right).$$

The center manager's problem is to find a booking limit combination to best balance the expected revenue and expected overage cost. In this study, we follow the idea of [21,22] to define the expected *net reward* as the expected revenue minus expected overage cost (the same concept is also used by [9] but called net profit therein). More precisely, the center manager chooses N_1, N_2, ..., and N_m to maximize the expected net reward

$$\sum_{x_1=1}^{N_1}\sum_{x_2=1}^{N_2}\cdots\sum_{x_m=1}^{N_m}\prod_{j\in J}\Pr(X_j = x_j)\pi(x_1, x_2, ..., x_m)$$

where

$$\pi(x_1, x_2, ..., x_m) = \sum_{j\in J} x_j P_j - \sum_{i\in I} f_i\left(\sum_{j\in J} V_{ij}x_j - K_i\right),$$

is the net reward when x_j consumers show up to do set j with no late cancellation.

2.2 Logistic Regression, Clustering, and Monte Carlo Simulation

To apply the aforementioned overbooking optimization model, the only issue remains is to estimate $\Pr(X_j = x_j)$, the probability for x_j consumers to show up for set j. One naïve way is to collect historical booking and cancellation data for a period of time. If there are q_1 late cancellations among q_2 bookings, one may use $r = \frac{q_1}{q_2}$ to be the late cancellation probability for each booking. The random variable X_j then follows the binomial distribution with N_j trials and success probability $1 - r$. However, as different consumers may have different show-up rates, finer estimation is possible.

To better estimate $\Pr(X_j = x_j)$, it is believed that consumers booking set j may have different late cancellation probability. For the kth booking for set j in a day, let r_{jk} be her/his probability of late cancellation. As r_{jk} may depend on many factors, once historical data are collected, one may first construct a logistic regression model to find out variables having significant correlation with late cancellation. For the regression model, the dependent variable is 1 if a booking is late canceled or 0 otherwise. One may either input independent variables according to her/his belief and domain knowledge, manually removing insignificant independent variables, or using variable selection methods like forward stepwise selection, backward stepwise selection, etc. To facilitate clustering, it is suggested to factorize all continuous independent variables.

As long as factors significantly correlated with late cancellations are identified, one may use those factors to cluster historical bookings. For example, if all consumers are either male or female, and gender is identified to be a significant factor, we may create two clusters containing bookings made by male and female consumers, respectively. If whether the examination is conducted in winter (which is defined as, say, November, December, and January) is also significant, each of the two clusters may be further split into two smaller clusters according to the examination season. To estimate the late cancellation probability of the kth booking for set j, one may then first find the cluster it belongs to according to its attributes and use the proportion of bookings with late cancellation in that cluster to be the late cancellation probability r_{jk}.

With r_{jk}s estimated for all N_j bookings requiring set j, $\Pr(X_j = x_j)$ may now be calculated. Theoretically, if N_j is small, this probability may be calculated by completely enumerating all the $\binom{N_j}{x_j}$ combinations for x_j out of N_j consumers to show up and summing up their probabilities. When N_j is large enough, however, this method is impractical, and we resort to the following Monte Carlo simulation with T iterations. In the tth iteration, we generate the number of consumers who do not cancel late by the following steps. Firstly, we decide whether a particular consumer show up by comparing a uniform random number $R \in [0, 1]$ and her/his late cancellation rate r_{jk}. The consumer is considering showing up without late cancellation if and only if $R \geq r_{jk}$. After doing this lottery for all the N_j consumers, we count the number of consumers who show up and denote it by $\hat{x}_j^t$. After T iterations are done, with $\{\hat{x}_j^1, \hat{x}_j^2, ..., \hat{x}_j^T\}$, we estimate $\Pr(X_j = x_j)$ as

$$\frac{\sum_{t=1}^{T} \mathbb{1}\{\hat{x}_j^t = x_j\}}{T}$$

for all $x_j = 0, 1, ..., N_j$, where

$$\mathbb{1}\{\hat{x}_j^t = x_j\} = \begin{cases} 1 \text{ if } \hat{x}_j^t = x_j \\ 0 \text{ otherwise} \end{cases}.$$

We keep increasing T until until $\Pr(X_j = x_j)$ converges.

2.3 Search for the Optimal Booking Limits

Combining the mathematical model and estimated probabilities obtained in the previous two subsections, one may now optimize the booking limits for all sets. It is natural to evaluate the performance of three booking policies: *no overbooking* (the NO policy) *overbooking with uniform probability* (the OU policy), and *overbooking with clustering* (the OC policy). The NO policy allows no overbooking. Under policy OU, overbooking is allowed while all bookings are assumed to have same late cancellation probability estimated as the proportion of late cancellations in all historical bookings. A search for optimal booking limits is then needed. As for the OC policy, the late cancellation probability of a booking is estimated as the proportion of late cancellations in the cluster it belongs to. The optimized booking limits may thus be different under the two overbooking policies.

To search for optimal booking limits, first a search space must be defined by determining the *maximum allowable overage* for each resource, which should be set to a level so that overage above this level is not acceptable. For resource i, let U_i be the maximum allowable overage of resource i, a constraint

$$\sum_{j \in J} V_{ij} N_j \leq K_i + U_i \tag{1}$$

is added into the stochastic mathematical model to ensure that the overage level can never exceed U_i. The search space is then defined as

$$S = \left\{ (N_1, ..., N_m) \middle| \sum_{j \in J} V_{ij} N_j \leq K_i + U_i \forall i \in I \right\}.$$

For the NO policy, U_i is set to 0 for all $i \in I$ to ensure no overage with 100% probability.

Depending on the number of sets and the magnitudes of booking limits, one may either do a complete search or a greedy search. A complete search considers all booking limit combinations $(N_1, ..., N_m)$ in the search space S, i.e., satisfying all these constraints. The expected net rewards of all feasible combinations are evaluated, and the combination with the highest expected net reward is considered optimal and reported. If m is too large or U_is are too large, a complete search

may not be finished in a reasonable amount of time. One may then resort to a greedy search. Starting from the combination $(N_1, N_2, ..., N_m) = (0, 0, ..., 0)$, i.e., the original booking limits, a greedy search runs several iterations. In each iteration, it evaluates the potential of each set by calculating the expected net reward resulted from increasing the booking limit of the set by one unit and keeping all other booking limits unchanged. The set with the highest potential is selected, the corresponding booking level is increased by 1, and then the next iteration starts. The greedy search stops when no increment brings in higher expected net reward or all increments lead to a combination outside the search space S. Note that the greedy search may not always obtain a truly optimal booking limit combination. When one applies it, one needs to study the optimality gap generated by the greedy search and modify the greedy search when necessary.

2.4 Discrete-event Simulation and Performance Evaluation

Once the optimal booking limit combination for a policy is obtained, we may evaluate it performance by calculating its expected net reward. As the real environment is actually uncertain, the calculation may be done by a discrete-event simulation. For the policy, we test its optimal booking limit combination by simulating bookings and late cancellations following the probability distributions estimated based on historical data. The simulation may be conducted in any time unit as one wants, say, day, week, month, or year. For example, suppose that one wants to evaluate the expected daily net rewards for a policy, she/he may generate bookings for a day following historical distributions. The first N_j^* bookings will be accepted, where N_j^* is the booking limit for set j in the optimal booking limit combination. Some of these accepted bookings will then be canceled late, where the occurrence of late cancellations also follows historical distributions. The realized revenue of each set and realized overage cost of each resource may then be combined to calculate the net reward for this day. The process may then be repeated multiple times, say, 50 times, for one to collect several daily net rewards. The average of these daily net rewards is then an estimate of the expected net reward brought by the optimal booking limit combination for this policy.

Using the discrete-event simulation described above, we may obtain the average net rewards and net reward distributions for the three booking policies. Their relative performance with respect to expected net rewards may then be compared, and statistical tests such as

Tucky HSD test may then be conducted to test the significance of differences. Note that while the optimal booking limit combinations of the NO and OU policies are obtained by assuming a uniform late cancellation probability, in the discrete-event simulation the late cancellation probabilities should be the real ones, which are estimated by the ones used in the OC policy.

Finally, beside expected net rewards, other performance indicators may be also important. To name a few, the expected number of consumers served is a better indicator of healthcare service accessibility, and the expected total overage is more related to service quality. As these quantities are required during the calculation of the expected net reward, these indicators may also be calculated through the discrete-event simulation process.

3 Results

To demonstrate how to develop an overbooking policy using the proposed way and evaluate its performance, we collaborate with a physical examination center in Taipei, Taiwan to conduct a case study.

3.1 Overview of the Physical Examination Center

Before our collaboration, the center does not adopt overbooking. The sum of booking limits for all sets, which is determined by center staffs using experience, is around 30. These spots are basically fully reserved everyday, and most of them are booked 180 days prior to the examination days, i.e., booked at the first day when reservation is open.

Appointment cancellations and rescheduling occur a lot everyday in the physical examination center. While most cancellations and rescheduling can be resolved by waiting for new appointments or moving later appointments to an earlier day to fill the vacancy, cancellations within seven days of the examination day are generally too late to have someone filling the spot. These cancellations are thus defined as late cancellations in this case study. In 2019, around 4% of bookings were canceled late, and late cancellations caused roughly 12,000,000 New Taiwan Dollar (NTD) losses in revenue and around 4% of examination opportunities wasted. Given that there are always many consumers waiting for examination, overbooking is considered.

The center serves three types of consumers: the general public, employees of companies having special contracts with the center, and government officials. In 2019, out of the total 12,390 bookings, 8,265 (66.7%), 2,966 (23.9%), and 1,159 (9.4%) consumers were of the three types, respectively. As the general public counts for the most bookings, and spots must be reserved to these companies and the government, in this case study we only focus on the general public, among which around 3.3% bookings were cancelled late.

3.2 Data Preprocessing and Exploratory Analysis

In 2019, in total 22 examination sets requiring eight resources were offered for consumers to self-select. The examination sets and their corresponding resources are in Table 13 in the appendix. Prices for the sets range roughly from 20,000 NTD to 100,000 NTD. Among the 22 sets, the top nine sets counted for more than 95% of the total bookings. To reduce computation complexity without sacrificing accuracy too much, we choose to consider only these nine sets. The resources that are under consideration is then limited to basic check, breast check, gastroenteroscope, magnetic resonance imaging (MRI), and diffusion weighted imaging (DWI). Furthermore, we aggregate the nine sets into four "aggregate sets" so that sets in the same aggregate set use the same resources. The sets, aggregate sets, and required resources are listed in Table 2. In the sequel, we call aggregate sets as set for simplicity. The picked sets and their corresponding booking percentages are put in the appendix.

Among the 8,265 bookings, some consumers rescheduled their original booking to a new examination date. In this case, a booking is split into two bookings, where the first one's cancellation date and the second one's booking date are both set to be the date of rescheduling. 326 new bookings are thus created. After a data cleansing process whose details are presented in the appendix, 160 bookings with unrecoverable missing values are removed. The number of remaining bookings is now 8,431, among which 8,046 bookings are for the top nine sets aforementioned. These 8,046 bookings form the final data set for the regression analysis. The proportion of late cancellation is 3.34% in the final data set, slightly higher than 3.3% in the original 8,265 bookings.

For each bookings, nine variables are recorded, including the consumer's gender and age, dates of booking, examination, and cancellation (if applicable), the chosen examination set, etc. These original independent variables are listed in the first part of Table 3. We then create five derived independent variables, listed in the second part of the table, using the original ones. The last part of the table contains the dependent variable, late cancellation, which is derived by comparing the examination and cancellation dates.

Table 2 Aggregate Sets, Sets, and Required Resources

Aggregate Set	Set(s)	Resource				
		basic	breast	gastroenteroscope	MRI	DWI
1	Male Basic Check Plus Colonoscopy	√		√		
2	Female Basic Check Above 40 Plus Colonoscopy Female Basic Check Below 40 Plus Colonoscopy	√	√	√		
3	Male Exquisite Anticarcinogenic (with MRI) Female Exquisite Anticarcinogenic (with MRI)	√			√	
4	Male Elite (with MRI + DWI) Female Elite (with MRI + DWI) Male Advanced Anticarcinogenic (with MRI + DWI) Female Advanced Anticarcinogenic (with MRI + DWI)	√			√	√

Table 3 Variables

Variable	Description	Data Type
Age	Age of the consumer	Numerical
Gender	Gender of the consumer	Categorical
Booking Date	The date when the consumer books an examination	Date
Booking Type	The way the consumer makes a reservation (in-person, online, or other)	Categorical
Examination Date	The date the consumer receives health examination	Date
Examination Set	The examination set chosen by the consumer	Categorical
Price	Price of the examination	Numerical
Group Size	Number of people booking together	Numerical
Cancellation Date	The date of cancellation (if applicable)	Date
Appointment Lead Time	Number of days between booking date and examination date	Numerical
Gastroscope	Whether the examination includes gastroscope or not	Categorical
Holiday Examination	1 if the examination is in a national holiday or 0 otherwise	Binary
Group Booking	1 if the group size is at least two or 0 otherwise	Binary
Last-minute Booking	1 if the appointment lead time is no greater than seven days or 0 otherwise	Binary
Late Cancellation	1 if canceled within seven days of the examination date or 0 otherwise	Binary

Before conducting the regression analysis, we try to find out important factors that are correlated with late cancellations through exploratory data analysis. To facilitate clustering of bookings, which will be conducted in later steps, various ways to divide numerical variables (such as age and price) and group date variables (such as examination date and booking date) are investigated. For example, from the raw data it is observed that January, February, August, November, and December have late cancellation probabilities higher than the overall average. We thus manually create several ways of grouping months, such as "November to February, July to August, others", "November to January, August, others", etc., and do regression with each way of grouping to see which one provides the best explanatory power. As the grouping "November to January, July to August, others" best explains late cancellations in our study, it is adopted in the final model. The grouping of other variables are also in a similar manner.

After some exploration, four variables are worth reporting, including age interval, examination season, group booking, and last-minute booking. The visualization is provided in Figure 1 with detailed numbers provided in Table 12 in the appendix. The plots and numbers regarding gender and examination sets are also provided just for reference.

The above process generates some managerial insights. First of all, consumers within the age interval 45 to 70 end up having lower late cancellation probability in average. This is not surprising, as middle-age people typically care about their health condition the most. Second, the average late cancellation probability is much higher in winter (November, December, and January) and summer (August, September) compared with the rest of the year. According to the staffs in the collaborating center, the huge number of events before New Year (both January 1 and the Chinese New Year) raises the late cancellation probability in winter. Similarly, events close to the end of summer vacation raise the probability in summer. Third, those booking with others have higher late cancellation probability in average than those booking individually. This can be due to the fact that when two or more consumers plan to go for physical examinations at the same day, they

may all cancel together whenever one of them cannot make it. Finally, last-minute bookings, though seldom occur, have much higher late cancellation probability than usual bookings. The staffs also have no idea about the reason. Regardless of fitting one's intuition or not, the exploratory analysis suggests us about variable creation and selection for regression model building.

3.3 Late Cancellation Probability Estimation

To identify significant factors that are correlated with late cancellation, we conduct logistic regression with late cancellation as the dependent variable. Three sets of independent variables are considered for our collaborating center. First, we consider those variables reported to be useful in literature (cf. Table 1) that we may collect, including age, gender, appointment lead time, examination season, and examination set. Continuous variables are factorized by categorizing bookings into interval following our exploratory data analysis. For the variable age, we create two intervals, namely between 45 and 70 or not. For appointment lead time, we create two levels, namely no greater than seven days (last-minute appointment) or not. For examination season, three levels are created, namely the winter season including November, December, and January, the summer season including August and September, and the off-peak season including the remaining seven months. Second, we input the three derived independent variables that seem to be effective in the exploratory analysis, including group booking, last-minute booking, and holiday examination (cf. Table 3. Whether these two sets of variables are statistically significant (with 0.05 as the threshold of p-value) is examined, and those insignificant variables are removed from the regression model. After this step, the remaining independent variables we have that are untested (such as price) are input as the third set of variables. A variable is kept in the model if it is statistically significant.

The above model building process results in the final regression model presented in Table 4. At the end, five factors are significant, including age interval, examination season, holiday examination, group booking, and last-minute booking. While intuitions obtained from exploratory data analysis (cf. Figure 1 and Table 12) are mostly confirmed, gender and examination sets turn out to have no significant correlation with late cancellations. According to the regression report, the risk of late cancellations is higher if a consumer is too young (below 45) or too old (above 70), the examination is in winter or summer (especially winter), the examination is in a national holiday, the booking is a group booking, and the booking is done in the last minute.

According to the categorical values of the five significant factors, theoretically each booking may be clustered into one of $2 \times 3 \times 2 \times 2 \times 2 = 48$ clusters. Nevertheless, because there are two few bookings that are in holidays or made in the last minute (cf. Table 12), these two factors are not used in clustering to avoid too sensitive estimation. Consequently, we obtain $2 \times 3 \times 2 = 12$ clusters in total. The clusters and the proportion of bookings having late cancellations in each cluster are shown in Table 5. Unsurprisingly, the cluster of winter examination, group booking, and age not within 45 and 70 has the highest proportion of late cancellation (6.07%). On the contrary, the cluster of off-peak examination, individual booking, and age within 45 and 70 has the lowest proportion (1.81%). The highest proportion is about 3.5 times higher than the lowest one.

The clustering result provides us finer estimation of the late cancellation probability of a booking. For each booking, we find the cluster it belongs to and use the historical late cancellation rate of the cluster as our estimated late cancellation probability. As an example, if a 38-year-old consumer makes an individual booking with no others in August, this booking will be clustered into cluster 8, and the estimated late cancellation probability will be 5.14%, rather than 3.34% as the overall historical late cancellation proportion.

It should be noted that, while in our study months should be categorized in to three groups (November to January, July to August, and others) to create the best model, it is by no means also the best for other studies. Different studies may have different research objectives, and customers in different regions may have different behaviors. Instead of simply using the ways of grouping adopted in our study, researchers are suggested to find their own ways to group months as well as other variables.

3.4 Booking Limit Optimization

As we introduce in Section 2.3, we consider three booking policies, no overbooking (NO), overbooking with uniform probability (OU), and overbooking with clustering (OC). Recall that the five resources considered in this case study are basic check, breast check, gastroenteroscope, MRI, and DWI. For the NO policy, which represents the current policy prior to this study, the ideal consumer level for the five resources are $K_1 = 36$, $K_2 = 36$, $K_3 = 24$, $K_4 = 4$, and $K_5 = 2$.

Both the NO and OU policies assume that all bookings have the same late cancellation probability, which is 3.34% in this study. Under the OC policy, however, the late cancellation probabilities are assumed to be different from cluster to cluster according to Table 5.
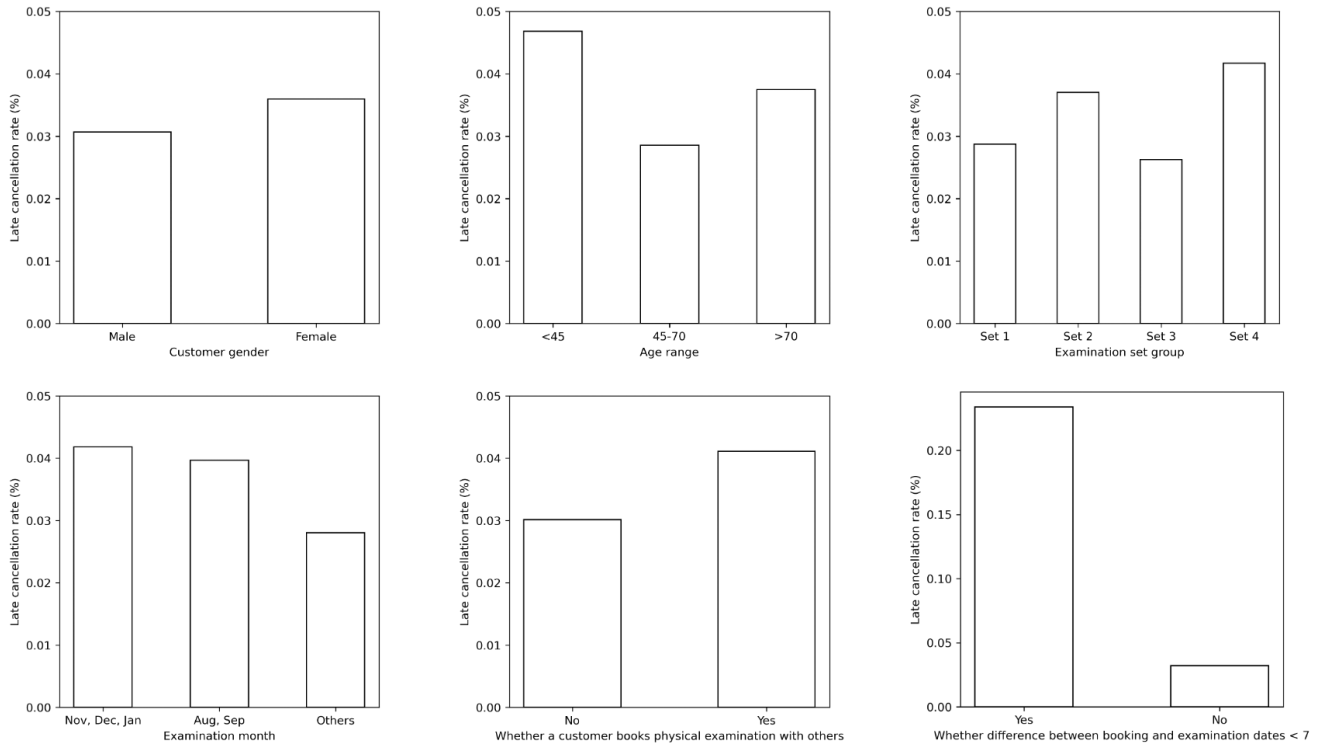
Fig. 1 Visualization of late cancellation rates

Table 4 Regression Report

Variables	Coefficient	Standard Error	z	p-value
Constant	−3.3203	0.112	−29.518	≈ 0
Age Interval (level: [45, 70])	−0.4303	0.116	−3.702	≈ 0
Examination Season (level: summer)	0.3711	0.135	2.754	0.006
Examination Season (level: winter)	0.4147	0.158	2.627	0.009
Holiday Examination (level: yes)	0.5702	0.311	1.836	0.066
Group Booking (level: yes)	0.2861	0.121	2.356	0.018
Lase-minute Booking (level: yes)	2.0522	0.325	6.321	≈ 0

Table 5 Clusters of Booking

Cluster	Examination Season	Group Booking	Age Interval: [45, 70]	Number of Bookings	Number of Late Cancellations	Late Cancellation Probability
1	Winter	Yes	Yes	583	20	3.43%
2	Winter	Yes	No	247	15	6.07%
3	Winter	No	Yes	842	31	3.68%
4	Winter	No	No	433	22	5.08%
5	Summer	Yes	Yes	236	10	4.24%
6	Summer	Yes	No	111	4	3.60%
7	Summer	No	Yes	578	19	3.29%
8	Summer	No	No	311	16	5.14%
9	Off-peak	Yes	Yes	835	32	3.83%
10	Off-peak	Yes	No	420	19	4.52%
11	Off-peak	No	Yes	2319	42	1.81%
12	Off-peak	No	No	1131	39	3.45%

As overbooking is not allowed under the NO policy, the maximum allowable overage $U_i = 0$ for all resource $i \in I$, and the search for an optimal booking limit combination is subject to the constraint in (1) with K_1 to K_5 as the right-hand-side values. By adopting the set-resource relationship recorded in Table 2, the five constraints may be listed as

$N_1 + N_2 + N_3 + N_4 \leq 36, N_2 \leq 36,$

$N_1 + N_2 \leq 24, N_3 + N_4 \leq 4,$ and $N_4 \leq 2.$

After discussions with center staffs, for the OU and OC policies, the maximum allowable overage of a resource is set to be 50% of the ideal consumer level of that resource, i.e., $U_i = \frac{1}{2}K_i$ for all $i \in I$. The right-hand-side values of these constraints therefore become 54, 54, 36, 6, and 3, respectively.

One may wonder why the maximum allowable overage rate 50% is much higher than 3.34%. To answer this, it should be noted that before the optimization is done, one has no way to know to what degree the original booking limits should be increased. While the proposed algorithm will search for optimal booking limits, one must first set a large enough search space for the algorithm not to miss the optimal booking limits. This is why a high maximum allowable overage rate is generally suggested. After all, setting $U_i = \frac{1}{2}K_i$ does not mean the optimal booking limits must result in a consumer level being equal to $\frac{3}{2}K_i$ for some resource i. The optimal booking limits will be determined by taking the overage costs into consideration. The higher the overage costs, the lower the optimal booking limits.

It remains to determine the overage cost function $f_i(z)$ for each resource $i \in I$ and overage level z. After discussions with center staffs, the marginal overage costs for the first ten units of overage are estimated by the collaborating center. In particular, the overage cost function, which is mainly based on overtime works from the perspective of the center, is set to be identical for all resources. It is estimated that each of the first three unit of overage incurs an overage cost of 2,500 NTD, each of the fourth to the six incurs a cost of 5,000 NTD, each of the seventh to the ninth incurs a cost of 10,000 NTD, and the tenth incurs a cost of 20,000 NTD. The accumulative overage cost of a resource is the sum of all marginal overage costs. For example, if for a resource the overage level is five, the total overage cost of that resource will be $2,500 \times 3 + 5,000 \times 2 = 17,500$ NTD. For overage levels over ten, center staffs find it too difficult to imagine and estimate. A second-order trend curve with zero intercept, $y = 18.802x + 682.73x^2$, is thus generated using the accumulative overage costs for the first ten units (cf. Figure 2). The accumulative overage cost

of eleven and twelve units of overage may then be estimated as 82,871 NTD and 98,539 NTD, respectively. It is clear that the spillover effect of overage exists as the marginal cost is increasing in the overage level.
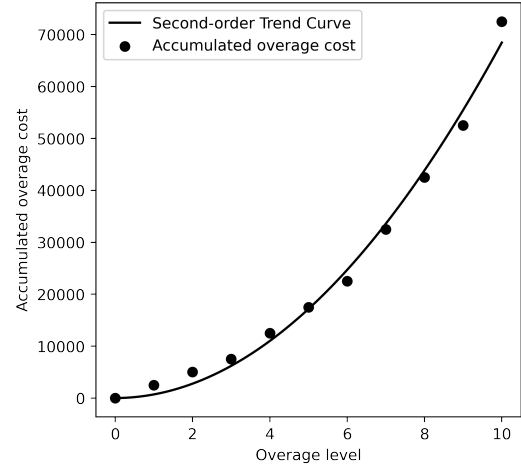

Fig. 2 Overage Costs and the Second-order Trend Curve

We are now ready to search for optimal booking limit combinations for the three booking policies. For this case study, the size of the search space is manageable, and thus complete search is adopted. After running through all feasible combinations, the search may return the one with the highest expected net reward for each policy. Note that as the late cancellation probabilities in different seasons are considered to be different in the OC policy, it is natural for the optimal combinations under the OC policy to differ in seasons. Table 6 lists the optimal combinations, where N_j^* is the booking limit of resource j in the optimal combination. It may be observed that the booking limits are higher under the overbooking policies. Moreover, under the OC policy, N_2^* is the lowest in the off-peak season. This is because that late cancellations are believed to occurs less likely in the off-peak season.

Table 6 Optimal Booking Limits under Different Policies

Policy	Season	N_1^*	N_2^*	N_3^*	N_4^*
NO	–	1	23	2	2
OU	–	1	35	3	3
OC	Winter	1	35	3	3
OC	Summer	1	35	3	3
OC	Off-peak	1	34	3	3

Note that a set of optimal booking limits must satisfy the resource constraint (1). For example, under the NO policy, the booking limits for the first two sets are

$N_1^* = 1$ and $N_2^* = 23$. The constraint for resource 3 ($N_1 + N_2 \leq 24$) makes it infeasible to accept more consumers for these two sets. Similarly, the constraint for resource 4 ($N_3 + N_4 \leq 4$) disallows the physical examination center to accept more consumers for the last two sets. This explains why the optimal booking limits only sum to 28 rather than 36 (the maximum among $K_i + U_i$ for all $i \in I$).

3.5 Performance Evaluation and Comparison

With the optimal booking limit combinations for the three policies, we apply the true late cancellation probabilities, which are estimated by the the ones used in the OC policy, to evaluate the performance of each policy. The performance indicator we consider here is weekly expected net reward, which is estimated through discrete-event simulation. For each policy, we simulate bookings and late cancellations for 50 weeks. For each policy in each season, the average weekly rewards (in NTD) is presented in Table 7, and the distribution of the 50 average weekly rewards is visualized in the box-plot in Figure 3.

Table 7 Average Simulated Weekly Net Reward

Policy	Weekly Net Reward		
	Summer	Winter	Off-peak
NO	4,903,168	4,892,003	4,943,680
OU	6,771,068	6,761,773	6,815,793
OC	6,781,929	6,784,008	6,809,758

From the simulation, we may see that overbooking indeed increases the average weekly net reward (as the NO-OU difference and NO-OC difference are both high), but whether clustering is beneficial for our collaborating center is unclear (as the difference between OU and OC is slim). Table 8, which contains the p-values reported by the Tukey HSD test about pairwise difference, confirms the observation. This somehow suggests that the probability estimation part does not generate significant benefit in this case study, which may be due to the fact that the late cancellation probabilities for all types of consumers are all low. Note that the p-value between OU and OC is the smallest in the winter season, in which the overall late cancellation probability is the highest among the three seasons. This suggests that probability estimation may still be beneficial in cases with higher overall late cancellation probability.

A final remark should be made regarding the significant improvement on the expected net reward. According to the above experiments, adopting the optimal

Table 8 Result of Tukey HSD test

Policy	p-value		
	Summer Peak	Winter Peak	Off-peak
NO vs. OU	0.001	0.001	0.001
NO vs. OC	0.001	0.001	0.001
OU vs. OC	0.562	0.175	0.824

overbooking policy may help the physical examination center to increase its expected net reward by around 38%. It is important to note that, however, this is due to the fact that in the center currently the estimated overage costs are relatively low (compared to the prices of typical examinations). Once the overage costs are increased (e.g., due to an increase in overtime wages or a higher need to ensure consumer satisfaction), both the optimal booking limits and increased expected net reward should decrease.

To see this, we repeat the discrete-event simulation in nine additional settings with the overage costs become two to ten times as much as the original ones. As the influence of costs is similar among the three seasons, we only demonstrate the analysis outcome in winter to save space. The impact of increasing the overage cost on the optimal booking limit combination and that on expected net reward are illustrated in Figures 4 and 5. As the NO policy does not overbook, the NO policy is not affected. On the contrary, both the OU and OC policies overbook less as the overage costs become larger. The expected net reward also decreases in the overage costs. While the result is intuitive, it also demonstrate the applicability of our proposed solution in different scenarios.

4 Discussion

Though the two overbooking policies are demonstrated to be effective in increasing expected net reward, it remains questionable how effective they may be in different scenarios. It is also important to ask whether they sacrifices too much in other perspectives. In this section, we conduct analysis with respect to overage cost estimation error, number of consumers served, and overage level to address these issues.

4.1 Impact of the Overage Cost Estimation Error

In practice, it can be hard in many cases for a physical examination manager to estimate the cost by evaluating the loss brought by overtime works and patient waiting in monetary values. If the overage cost is greatly underestimated, overbooking may actually bring in a
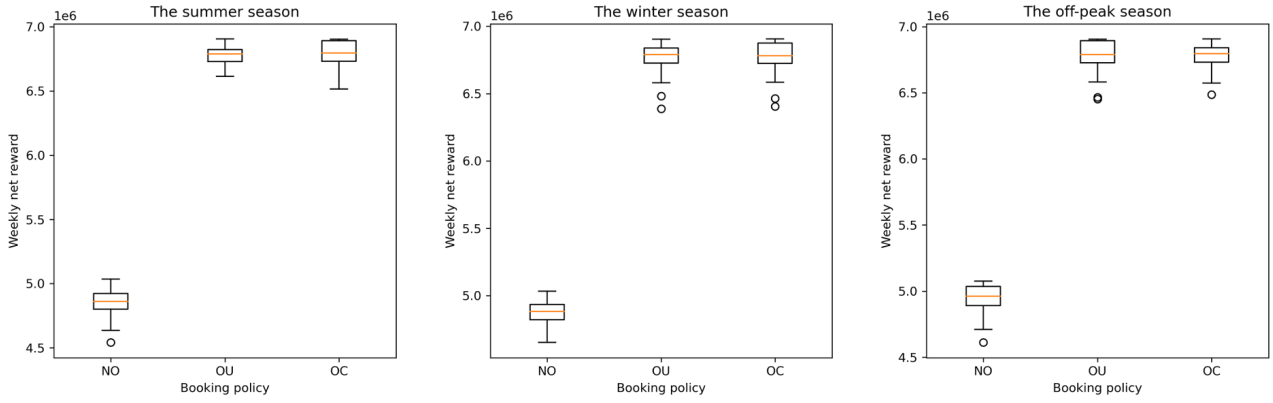
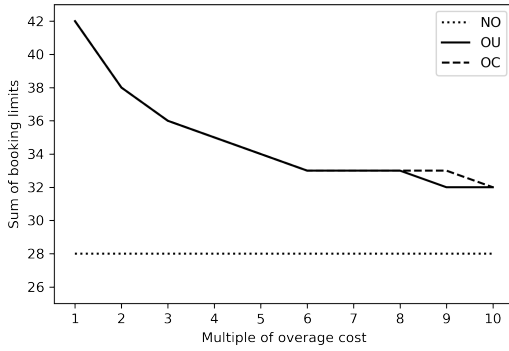**Fig. 3** Boxplot of Simulated Fifty Weekly Net Rewards under Different Booking Policies

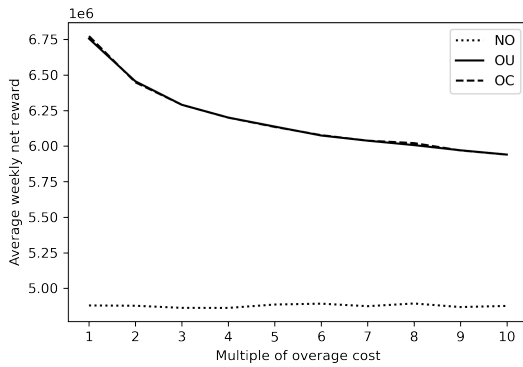**Fig. 4** Impact of Overage Cost on the Optimal Booking Limits in Winter

Fig. 5 Impact of Overage Cost on the Expected Net Rewards in Winter

estimated one, the expected net rewards under the two overbooking policies are still higher than that under no overbooking. Overbooking is thus quite applicable.
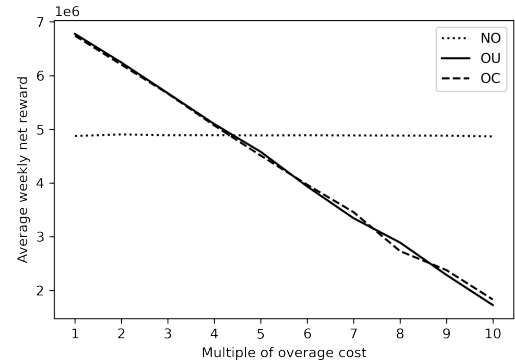

Fig. 6 Expected Net Reward under Overbooking when Underestimating Overage Cost

4.2 More performance indicators

We next investigate the number of customers served under the three policies. As shown in Figure 7, though the number of consumers served is random under the two overbooking policies, such a number is consistently higher than that under no overbooking. In other words, the accessibility of healthcare resources is improved by overbooking as resource wastes are reduced in average.

While overbooking is helpful in enhancing expected net reward and healthcare resource accessibility, it unfortunately creates overage, which results in patient waiting and overtime works. Even if the formula of net reward already include overage cost, it is still important to put an eye on the degree of overage. For each
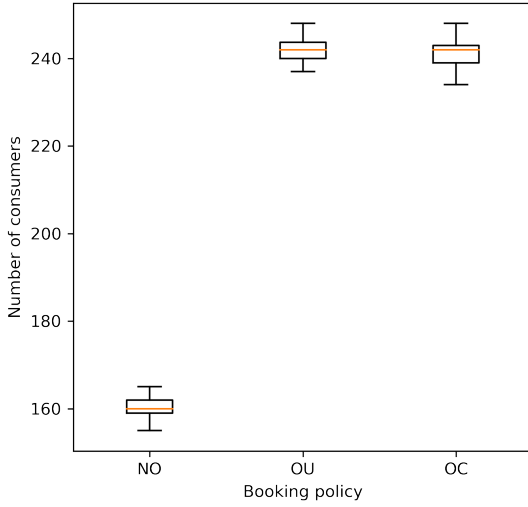
lower expected net reward. It is good if the benefit of overbooking is not too sensitive to the precision of overage cost estimation. Figure 6 depicts the result of a sensitivity analysis, in which the overage cost is actually multiple times of the one estimated by the decision maker adopting overbooking. For our collaborating center, even if the actual overage cost is four times of the

Fig. 7 Number of Consumers Served per Week in Winter

resource in a period of time, let the *overage percentage* be the overage level divided by the number of consumers served. According to Figure 8, which illustrates the distribution of the 50 weekly overage percentages under the OC policy in winter produced by the discrete-event simulation, in average roughly 10–11% of consumers served are considered overage. Considering the fact that the increase on expected reward is about 38% (cf. Table 7), the benefit seems to be large enough to compensate the detriment.
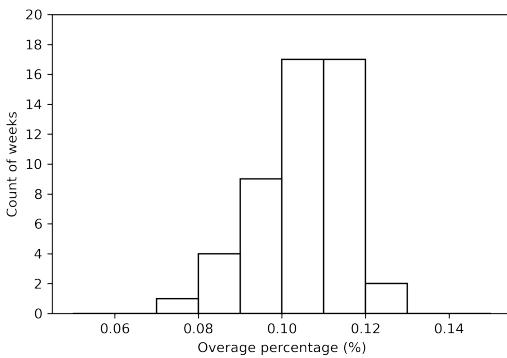
Fig. 8 Overage percentages under the OC policy in Winter

In many case, increasing the expected net reward means to overbook more, which means creating more overage in expectation. Facing the trade-off between expected net reward and expected overage percentage, the concept of Pareto frontier may help. In Figure 9, each circle is plotted according to the expected net re-

ward and expected overage percentage resulted from a booking limit combination of the OC policy in winter. The optimal booking limit combination, which brings in the highest expected net reward, is plotted as the only solid circle. While this combination is the best along the dimension of expected net reward, it is not as good as many other combinations along the dimension of expected overage percentage. If the expected overage percentage is considered too high to be acceptable, all the combinations that have been examined during the search process may serve as alternatives. Suppose that there is a maximum allowable expected overage percentage, say, 7%, one may turn to choose the combination plotted as the only solid square, which generates the highest expected net reward than all combinations having low enough expected overage percentage.
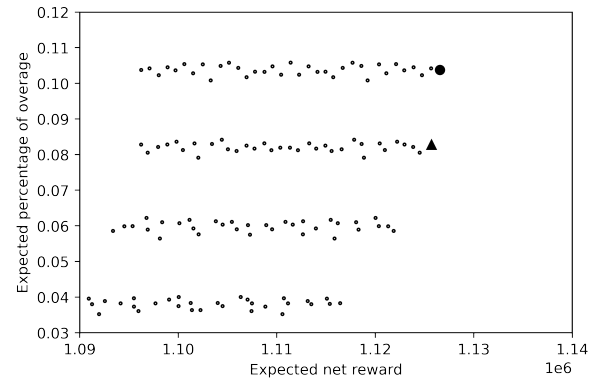

Fig. 9 Overage percentages under the OC policy in Winter

In general, a combination is called Pareto-efficient if there is no other combination having higher expected net reward and lower expected overage percentage than it. Our proposed solution may report all Pareto-efficient combinations identified during the search process to a physical examination manager in practice. A booking limit combination may then be selected according to the manager's decision criterion. The performance indicator expected overage percentage may also be substituted by expected overage level, expected overage cost, etc., to fit the need of a practitioner.

5 Conclusion

In this study, overbooking for physical examination considering the relationship between examination sets and resources is studied. A stochastic mathematical programming model is formulated for expected net reward optimization, a way of late cancellation probability estimation is developed with logistic regression, Monte

Carlo simulation, and clustering, and a way of performance evaluation is constructed through discrete-event simulation. With the help of a collaborating physical examination center in Taiwan, we conduct a case study to demonstrate the benefit of applying the proposed methods. In short, overbooking does help under the circumstance that late cancellation causes great loss. The higher the cancellation rate, the greater improvement on the expected reward.

This study is certainly limited and may be improved in several directions. First, probability estimation may be improved by using more relevant features. For example, an owner of a physical examination center in many cases also own clinics or hospitals. It may be promising to use a consumer's records in the clinic/hospital to estimate its late cancellation probability. Second, while the search space of booking limit combinations is small in this study, it may be quite big in general and calls for the need of a heuristic search. Developing an effective and efficient heuristic search algorithm with rigorous performance evaluation can also be valuable. In particular, as an overbooking problem is to determine the capacity level facing uncertain demand, it shares the same characteristics with the well-known newsvendor model in the operations management literature [30, 34].[1] While the set-resource relationship, integer constraints, and nonlinearity make the closed-form critical ratio formula cannot be directly applied, when a good heuristic algorithm is needed, algorithms solving variants of newsvendor problems are certainly great references. Finally, while in this study we do not consider interventions to reduce late cancellation probabilities, integrating overbooking with costly interventions may generate fruitful results.

List of abbreviations

1. NTD: New Taiwan Dollar.
2. MRI: Magnetic Resonance Imaging.
3. DWI: Diffusion Weighted Imaging.
4. CT: Computed Tomography.
5. NO policy: the No-Overbooking policy.
6. OU policy: the Overbooking policy.
7. OC policy: the Overbooking with Probability Estimation by Clustering policy.
8. Tukey HSD test: Tukey Honestly Significant Difference test.

---

[1] We thank an anonymous reviewer for pointing this out.

References

1. Ahmadi-Javid, A., Jalali, Z., Klassen, K.J.: Outpatient appointment systems in healthcare: A review of optimization studies. European Journal of Operational Research **258**(1), 3–34 (2017)
2. Bean, A.B., Talaga, J.: Predicting appointment breaking. Marketing Health Services **15**(1), 29–34 (1995)

3. Chen, Y., Kuo, Y., Fan, P., Balasubramanian, H.: Appointment overbooking with different time slot structures. Computers & Industrial Engineering **124**, 237–248 (2018)

4. Chen, Y., Kuo, Y.H., Balasubramanian, H., Wen, C.: Using simulation to examine appointment overbooking schemes for a medical imaging center. 2015 Winter Simulation Conference (WSC). IEEE pp. 1307–1318 (2015)

5. Chesanow, N.: Can't stay on schedule? here's a solution. Medical Economics **73**(21), 174–177 (1996)

6. Collins, J., Santamaria, N., Clayton, L.: Why outpatients fail to attend their scheduled appointments: a prospective comparison of differences between attenders and non-attenders. Australian Health Review **26**(1), 52–63 (2003)

7. D.Gupta, Denton, B.: Appointment scheduling in health care: Challenges and opportunities. IIE Transactions **40**, 800–819 (2008)

8. Dreiher, J., Froimovici, M., Bibi, Y., Vardy, D.A., Cicurel, A., Cohen, A.D.: Nonattendance in obstetrics and gynecology patients. Gynecologic and Obstetric Investigation **66**(1), 40–43 (2008)

9. Feldman, J., Liu, N., Topaloglu, H., Ziya, S.: Appointment scheduling under patient preference and no-show behavior. Operations Research **62**(4), 794–811 (2014)

10. Grunebaum, M., P. Luber, M.C., Leon, A.C., Olfson, M., Portera, L.: Predictors of missed appointments for psychiatric consultations in a primary care clinic. Psychiatric Services pp. 848–852 (1996)

11. Hassin, R., Mendel, S.: Scheduling arrivals to queues: A single-server model with no-shows. Management Science **54**(3), 565–572 (2008)

12. Ho, C.J., Lau, H.S.: Minimizing total cost in scheduling outpatient appointments. Management science **38**(12), 1750–1764 (1992)

13. Johnson, B.J., Mold, J.W., Pontious, J.M.: Reduction and management of no-shows by family medicine residency practice exemplars. The Annals of Family Medicine **5**(6), 534–539 (2007)

14. Kaplan-Lewis, E., Percac-Lima, S.: No-show to primary care appointments: why patients do not come. Journal of Primary Care & Community Health **4**(4), 251–255 (2013)

15. Kim, S., Giachetti, R.E.: A stochastic mathematical appointment overbooking model for healthcare providers to improve profits. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans **36**(6), 1211–1219 (2006)

16. Kuo, Y.H., Balasubramanian, H., Chen, Y.: Medical appointment overbooking and optimal scheduling: tradeoffs between schedule efficiency and accessibility to service. Flexible Services and Manufacturing Journal **32**, 72–101 (2020)

17. Lacy, N.L., Paulman, A., Reuter, M.D., Lovejoy, B.: Why we don't come: patient perceptions on no-shows. The Annals of Family Medicine **2**(6), 541–545 (2004)

18. LaGanga, L.R., Lawrence, S.R.: Clinic overbooking to improve patient access and increase provider productivity. Decision Sciences **38**(2), 251–276 (2007)

19. Lasser, K.E., Mintzer, I.L., Lambert, A., Cabral, H., Bor, D.H.: Missed appointment rates in primary care: the importance of site of care. Journal of Health Care for the Poor and Underserved **16**(3), 475–486 (2005)

20. Lee, V.J., Earnest, A., Chen, M.I., Krishnan, B.: 2005. Predictors of failed attendances in a multi-specialty outpatient centre using electronic databases **5**(1), 51–59 (BMC Health Services Research)

21. Liu, N.: Optimal choice for appointment scheduling window under patient no-show behavior. Production and Operations Management **25**(1), 128–142 (2016)

22. Liu, N., Ziya, S.: Panel size and overbooking decisions for appointment-based services under patient no-shows. Production and Operations Management **23**(11), 2209–2223 (2014)

23. Liu, N., Ziya, S., Kulkarni, V.G.: Dynamic scheduling of outpatient appointments under patient no-shows and cancellations. Manufacturing and Service Operations Management **12**, 347–364 (2010)

24. McGill, J.I., Ryzin, G.J.V.: Revenue management: research overview and prospects. Transportation Science **33**(2), 233–256 (1999)

25. Mieloszyk, R.J., Rosenbaum, J.I., Hall, C.S., Hippe, D.S., Gunn, M.L., Bhargava, P.: Environmental factors predictive of no-show visits in radiology: Observations of three million outpatient imaging visits over 16 years. Journal of the American College of Radiology **16**(4), 554–559 (2019)

26. Murray, M., Berwick, D.M.: Advanced access: reducing waiting and delays in primary care. Jama **289**(8), 1035–1040 (2003)

27. Muthuraman, K., Lawley, M.: A stochastic overbooking model for outpatient clinical scheduling with no-shows. IIE Transactions **40**(9), 820–837 (2008)

28. Parente, C.A., Salvatore, D., Gallo, G.M., Cipollini, F.: Using overbooking to manage no-shows in an italian healthcare center. BMC Health Services Research **18**(1), 185–196 (2018)

29. Perron, N.J., M-D. Dao, M.P.K., Miserez, V., Chuard, C., Calmy, A., Gaspoz, J.M.: 2010. Reduction of missed appointments at an urban primary care clinic: a randomised controlled study **11**(1), 79–86 (BMC Family practice)

30. Qin, Y., Wang, R., Vakharia, A., Chen, Y., Seref, M.: The newsvendor problem: Review and directions for future research. European Journal of Operational Research **213**(2), 361–374 (2011)

31. Rothstein, M.: Hotel overbooking as a markovian sequential decision process. Decision Sciences **5**(3), 389–404 (1974)

32. Rust, C.T., Gallups, N.H., Clark, W.S., Jones, D.S., Wilcox, W.D.: Patient appointment failures in pediatric resident continuity clinics. Archives of Pediatrics & Adolescent Medicine **149**(6), 693–695 (1995)

33. Shonick, W., Klein, B.W.: An approach to reducing the adverse effects of broken appointments in primary care systems: development of a decision rule based on estimated conditional probabilities. Medical Care pp. 419–429 (1977)

34. Turken, N., Tan, Y., Vakharia, A.J., Wang, L., Wang, R., Yenipazarli, A.: The multi-product newsvendor problem: Review, extensions, and directions for future research. In: T.M. Choi (ed.) Handbook of Newsvendor Problems: Models, Extensions and Applications, pp. 3–39. Springer New York, New York, NY (2012)

35. Yan, C., Huang, G.G.Q., Kuo, Y.H., Tang, J.: Dynamic appointment scheduling for outpatient clinics with multiple physicians and patient choice. Journal of Management Science and Engineering (2021)

Appendix A: Detailed Data Cleansing Process

Because the impact of appointment cancellation and rescheduling request is similar, we view a rescheduling request as a cancellation combined with a new appointment. Therefore, for a booking with N bookings of changing examination dates, we break them down into

$N + 1$ new bookings, and for the second to last new bookings, the booking date is the cancellation date of the previous booking.

As an example, consider a booking made on February 3 asking for an examination on August 8. On April 7, the consumer asked for rescheduling to September 3. On June 2, the consumer further asked for rescheduling to October 15. The two rescheduling requests make the original booking be split into three bookings shown in Table 9.

Table 9 Bookings after splitting

Booking Date	Examination Date	Cancellation Date
2019-2-3	2019-8-5	2019-4-7
2019-4-7	2019-9-3	2019-6-2
2019-6-2	2019-10-15	N/A

We make some judging criteria of labeling rescheduling requests after discussing with physical examination center. First, Date exchanging between two consumers will not be viewed as changing date since the request will not cause losing on reward. Second, date changing caused by unpredictable factors, such as natural disasters or sudden closing of physical examination center, will not be viewed as rescheduling since that could not be controlled by the consumer. Third, since every set requires different resources, and changing sets will lead to additional administration costs, changing examination sets will be viewed as changing. However, the request of changing a single examination item will not be counted into changing.

There are some missing data in the raw data set. In very few cases, the date when consumers requested to change the examination dates are missing due to the human error of staffs. In these cases, we consider the average of the original booking date and the new examination date as the booking date of the new booking. After the process, we recover most missing data, but 160 bookings are still unrecoverable.

After cleansing and splitting, there are 8,431 bookings, roughly 2% more than the original number of citizen bookings in the raw data. As mentioned, we only consider the top nine sets in this study. Thus, the number of bookings further decreases to 8,046. Furthermore, the late cancellation rate is 3.3% in the original data and 3.34% in the cleansed data.

Appendix B: Supplemental Data

Table 10 lists the variables that are mentioned in other papers saying to have influence on the late cancellation rate but not used in this study.

Table 11 provides a list of the top nine examination sets and their corresponding booking percentages among all sets.

Table 12 provides a list of the numbers of bookings, numbers of late cancellations, and late cancellation probabilities under different clustering categories.

Table 13 provides a list of all the examination sets and their corresponding valuable resources.

Table 10 Factors considered in the Literature

	[6]	[8]	[10]	[13]	[14]	[17]	[20]	[28]	[32]
previous failure to attend	v						v	v	
traffic convenience of the hospital	v						v		
patient's race		v			v		v		
patient distress				v					
patient opinion concering referal				v					
patient's health insurance type					v	v		v	v
patient continuity of care					v				
emotional barrier						v			
concern about disrespect by the hospital						v			
lack of understanding of the scheduling system						v			
provision of cell phone number							v		
type of patient (new/old)								v	
booking confirmation/reminder service								v	
type of booking								v	
day of week								v	
time of day								v	
weather								v	

Table 11 Top Nine Examination Sets

Set name	Bookings	Percentage
Male Basic Check Plus Colonoscopy	3068	36.4%
Female Basic Check Above 40 Plus Colonoscopy	2754	32.7%
Female Basic Check Below 40 Plus Colonoscopy	760	9.0%
Male Exquisite Anticarcinogenic (with MRI)	315	3.7%
Female Exquisite Anticarcinogenic (with MRI)	333	3.9%
Male Elite (with MRI and DWI)	387	4.6%
Female Elite (with MRI and DWI)	233	2.8%
Male Advanced Anticarcinogenic (with MRI and DWI)	108	1.3%
Female Advanced Anticarcinogenic (with MRI and DWI)	88	1.0%
Others	385	4.6%
Total	8431	100.0%

Table 12 Late Cancellation Probability in Categories

Variable	Category	Bookings	Late Cancellations	Late Cancellation Probability
Gender	Male	3,875	119	3.1%
	Female	4,171	150	3.6%
Examination Set	Set 1	3,068	88	2.9%
	Set 2	3,514	130	3.7%
	Set 3	648	17	2.6%
	Set 4	816	34	4.2%
Age Interval	<45	1,666	78	4.7%
	45-70	5,393	154	2.9%
	>70	987	37	3.7%
Group Booking	No	5,614	169	3.0%
	Yes	2,432	100	4.1%
Last-minute Booking	Yes	47	11	23.0%
	No	7,999	258	3.2%
Season	Winter	2,105	88	4.2%
	Summer	1,236	49	4.0%
	Others	4,705	132	2.8%
Total		8,046	269	3.3%

Table 13 Examination Sets and required Resources

Examination Set	Description	Resources
Male Basic Check	The only difference between these two sets is the depth of the enteroscopy check. (sigmoid colon or the whole large intestine)	N/A
Male Basic Check Plus Colonoscopy		Gastroenteroscope
Female Basic Check Below 40	The differences between ages is the breast check. Breast ultrasound for female below 40 while mammography for the other.	Breast
Female Basic Check Above 40		Breast
Female Basic Check Below 40 Plus Colonoscopy		Breast + Gastroenteroscope
Female Basic Check Above 40 Plus Colonoscopy		Breast + Gastroenteroscope
Male Exquisite Anticarcinogenic (with MRI)	The differences between Exquisite and Advanced Anticarcinogenic is the magnetic resonance imaging (MRI). They both include normal MRI. While the diffusion weighted imaging (DWI) is added into the second one, which can help distinguish malignant tumor from benign tumor.	MRI
Female Exquisite Anticarcinogenic (with MRI)		MRI
Male Advanced Anticarcinogenic (with MRI + DWI)		MRI + DWI
Female Advanced Anticarcinogenic (with MRI + DWI)		MRI + DWI
Male Elite (with MRI + DWI)	The Elite is the Advanced Anticarcinogenic plus the coronary angiogram.	MRI + DWI
Female Elite (with MRI + DWI)		MRI + DWI
Male Exquisite Cardiac	The Advanced Cardiovascular is the Exquisite Cardiac plus head MRI and cardia echo.	Cardia Echo
Female Exquisite Cardiac		Cardia Echo
Male Advanced Cardiovascular		Cardia Echo + Cerebrum
Female Advanced Cardiovascular		Cardia Echo + Cerebrum
Male Gastroenteroscope	The six examination set is mainly for the government employee. They can choose one from gastroenteroscope, computed tomography coronary angiography (CTCA) or breast computed tomography (CT).	Gastroenteroscope
Male Computed Tomography Coronary Angiography		Cardia Echo
Male Breast Computed Tomography		CT
Female Gastroenteroscope		Gastroenteroscope
Female Computed Tomography Coronary Angiography		Cardia Echo
Female Breast Computed Tomography		CT