IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 2, 2019, accepted August 21, 2019, date of publication August 26, 2019, date of current version September 9, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2937687

Optimal Staffing Policy in Commercial Banks Under Seasonal Demand Variation

YUANQIAO CHEN[1,2], XIAOQIANG ZHANG[1], BOMIN BIAN[3], AND HAOHUA LI[1]

[1]School of Management and Engineering, Nanjing University, Nanjing 210093, China
[2]Urban Finance Research Institute, Industrial and Commercial Bank of China, Beijing 100140, China
[3]Institute of Systems Engineering, College of Management and Economics, Tianjin University, Tianjin 300072, China

Corresponding author: Haohua Li (howardlee_1990@hotmail.com)

This work was supported by the National Natural Science Foundation of China under Grant 71720107001 and Grant U1811462.

ABSTRACT According to evidence from research and practice, the seasonal demand variation exists in the banking sector and significantly affects the commercial bank's operation. Thus, the bank needs to adopt a seasonal staffing policy to ensure that the human resources can match the customer demands in different seasons. Otherwise, a shortage or surplus of human resource will occur and negatively affect bank operations. In this paper, we develop a seasonal staffing method to help the bank find the optimal staffing policy under seasonality. To capture the characteristics of bank operations, we model the service systems of n branches in a bank as a n-dimensional M/M/c/N queueing system with balking and reneging. Then a profit maximizing model based on the queueing system is constructed, and it is further simplified through linearization so that the model can be solved in a short time period. In addition, we conduct a set of numerical experiments that not only prove the superiority of our method compared with the traditional methods, but also explore the effects of some key factors on the optimal seasonal policy.

INDEX TERMS Banking, optimization, staffing, seasonal demand variation, queueing system.

I. INTRODUCTION

It is undoubtedly crucial for service institutions to enable their resources for services to match the customer demands. And when customer demands change, the resource allocation should be adjusted accordingly. Otherwise, shortage or surplus of resources will occur and cause a loss to the institution.

As an important type of financial service institution, the commercial bank is no exception to this rule. But in practice, to adjust resource allocation effectively as well as efficiently is not an easy task for the bank manager. In especial, the problem of human resource allocation (i.e., staffing) under demand variation troubles many bank managers because human resource management is one of the most important but challenging pieces of work in banking sector.

This paper focus on the issue with respect to making staffing policy under a common type of customer demand variation faced with commercial banks — the seasonal demand variation. The source of this seasonality is in the economy and society. As a key hub in economy, banking sector is affected any many economic factors, including seasonality.

The associate editor coordinating the review of this article and approving it for publication was Rajesh Kumar.

Seasonal demand variation faced with the commercial bank is mainly reflected by seasonal variation of customer flow, which exists in both counter service channels and online service channels. In this paper, we focus on the former since it highly relies on human resource while the latter doesn't. Plenty of evidence (see websites [1]–[3] for examples) shows significant demand variation between peak season and off-peak season in most of commercial banks. A group of investigations done by us in some Chinese banks last year show that customer flow in some branches in peak season could reach 2 to 3 times as high as that in off-peak season. The seasonal demand variation is especially evident in those branches located in rural areas mainly because of the seasonal nature of agriculture.

Under seasonal demand variation, some bank managers try to adopt a seasonal staffing policy: arranging a group of permanent servers in each branch, and sending additional staff members from the head office to each branch to act as a reinforcement in peak season. This raises two questions: (1) how many permanent servers should be arranged in every branch throughout the year? (2) how many additional staff members should be dispatched to each branch in peak season? Due to a lack of in-depth study on seasonal staffing problem, these bank managers can only decide on the numbers of permanent servers and additional staff based on a relatively rough estimation of the customer demands in different seasons. Then either overestimation or underestimation is likely to happen. The former results in overstaffing and further causes human resource waste, while the latter leads to understaffing and further breeds customer churn. Both the two situations will obviously cause loss and damage to the bank.

Aiming at helping the commercial bank avoid such loss and damage, we construct a theoretical profit maximizing model that can help the bank quantitatively optimize the staffing policy under seasonal demand variation. Based on existing research and practice, we model the counter service system in each branch in the bank as a limited-capacity M/M/c/N queueing system with balking (the customer may not join the line when he arrives to find he has to wait in line for service) and reneging (the customer may choose to leave while waiting).

And to ensure generality and practicality, we consider n (n can be any positive integer) branches in a commercial bank by including n queuing systems in our model. Usually, the commercial bank has more than one branch. Since the bank's human resource is limited, it is hard to achieve the local optimal human resource allocation in every branch simultaneously. Instead, the bank manager should seek for a global optimal assignment for all the branches. Obviously, the single-system model used by many other existing studies is not capable of finding the global optimal solution under a multi-system and resource-limited framework. The n-dimensional model, by contrast, is suitable for dealing with such issue because it completely fits into the multi-system framework. Therefore, compared to the model with single system, the n-dimensional model is more practical and general for the bank with more than one branch.

Based on the n-dimensional queueing model, we build a profit maximizing model that leads to an optimal solution to the problem of seasonal staffing. Thus, we obtain a model-based method of seasonal staffing. It should be noted that our method could also be spread to other service sectors faced with seasonality as long as some formulae and parameters are adjusted according to the specific circumstance of the service institution.

A set of numerical experiments are conducted after theoretical modelling to test the advantage of our model and explore each parameter's effects on our model. Parameter setting in the experiments is based on a set of data collected by surveys, which enables the experiments to be closed to the actual situation of the commercial bank so as to get practically meaningful results.

The rest of this paper is organized as follows. In Section II, a literature review is provided to further clarify academic problems of this research. Section III explicitly explains the theoretical model, including its construction, simplification and application. Section IV gives a set of numerical experiments based on actual data to test the advantage of our model-based method and further explore the effects of parameters. Finally, we provide a summary and future outlook in Section V.

II. LITERATURE REVIEW

The problem of staffing is recognized by both academics and practitioners as one of the key management problems in various industries, especially in those with high dependence on human resource, such as banking sector. It is proved by a few theoretical and empirical research that staffing optimization has significant positive impact on the bank's performance [4]–[6]. And improper staffing is detrimental to the bank's operation. Overstaffing results in human resource waste [7], [8], while understaffing leads to customer churn [9]–[11]. In essence, improper allocation of any kind of resource will weaken the supporting role of the resource [12]–[14].

It is a challenge with respect to staffing that changing economic and business environment make it necessary for the manager to adjust personnel dynamically [15]. Thus, research and practice of staffing should be combined with the trend of external environment. The seasonal demand variation is a common type of trend in the commercial bank's business environment. According to literature [16], [17], the formulation of seasonal demand variation faced with the bank can be outlined by Fig.1:

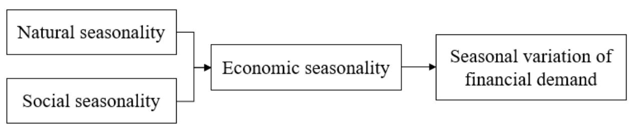


FIGURE 1. Formulation of seasonal demand variation faced by the bank.

Here, the sources of natural and social seasonality involve basic laws in sociology and economics. A lot of evidence proves the significance of the seasonality in financial market [18]–[20], indicating the effectiveness of the above transmission routes of seasonality.

However, there is little research on the seasonal staffing method for the bank to cope with significant seasonal demand variation. As we have discussed in Section I, this academic gap leads to unreasonable staffing policy-making in practice. Therefore, it is of great academic and practical significance to deeply explore the optimization of seasonal staffing policy under seasonal demand variation.

In order to do such research, we use the queueing theory to build a model of the bank's counter service system. Actually, there is an abundance of existing research using the queuing theory to model various service systems so as to further discuss the problem of resource allocation in the service institutions, such as literature [21]–[23]. And there is also research using the queueing theory to study the problem of task allocation, such as literature [24].

On staffing issues, it is showed by Rolfe [25] and Dyer & Proll [26] that a simple M/M/c queuing system can be used to get an optimal static staffing plan analytically. But M/M/c queuing system is obviously too simplistic to capture enough characteristics of the bank's service system. Ancker & Gafarian [27], [28] argue that there are likely to be balking and reneging phenomena in the actual queueing system. And they deduce the probability distributions of the queue length and waiting time in the M/M/1/N queueing system with balking and reneging. Nearly thirty years later, Abou-El-Ata & Hariri [29] give the probability distributions of the queue length and waiting time in the M/M/c/N queueing system with balking and reneging, which are applied to our modelling in this paper. Additionally, there are many scholars like [30]–[32] deepening and expanding the theory of the queueing system with balking and reneging.

Queueing theory is also widely used in studies on the call center, a very useful service network that provides telephoned-based services. The call center has become a fertile academic field that has attracted much attention with regard to forecasting, capacity planning, queueing, and personnel scheduling [33], [34]. There are big differences between the call center model and ours, although they both concern a dynamic and flexible policy to cope with changing customer demands [35].

Essentially, the operating modes of a telephone-based call center system and the bank's counter service system are distinct, mainly reflected by two big differences of the models. Firstly, while reducing the waiting time is one of the major objectives of the call center model, we do not take the waiting time into consideration in this research. The adverse effect of the customers' waiting is not ignored but is reflected by balking and reneging instead of waiting time.

Secondly, the call center model focuses on a relatively frequent staffing in which the costs mainly come from personnel shifting within the system as well as other operational activities, while our study concerns a far less frequent seasonal staffing between a head office and n different systems (they might be so distant from each other that frequent staffing is unfeasible), indicating that personnel shifting is not the major source of the staffing-related costs in our model. Instead, we mainly consider the salary which is recognized as one of the most important factors in human resource management. An employee's salary is related to his performance that is measured by his contribution to the bank. It is proved by lots of research that a reasonable mechanism of performance-related pay can significantly motivate the employees to work harder [36], therefore significantly promoting institution's operation efficiency [37], quality of service or production [38], [39], and even innovation [40]. Thus, by considering the performance-related pay in the cost function, we make our model link to the bank's operation more closely.

In addition, we have two innovations compared to existing research on the application of queueing theory in services. Firstly, unlike most of the existing research that just considers a single queueing system (such as [41]–[43]), we construct a n-dimensional M/M/c/N queueing model with balking and reneging (see Section I for the advantage of the n-dimensional model). Secondly, unlike lots of existing research that focus on cost reduction in the service system (such as [22], [23]), we build a profit maximizing model to fully embody the commercial nature of the bank.

Through considering many practical factors, we establish a nonlinear integer programming model. Inspired by the theory of linearization technique for solving discrete problem [44], [45], we find a way to linearize our model to make it easier to solve. The linearized model can be solved by the special branch and bound method [46]–[48].

Inspired by some literature on operations research (such as [49], [50]), we do a group of numerical experiments to test the effects of model parameters, and show the superiority of the seasonal staffing method based on the model. The results are consistent with the principle that a better match between customer demands and supporting resource can significantly enhance the bank's operation efficiency and bring more profits [4]–[6]. Besides, our numerical results reveal the importance of integrating human resources and physical resources properly, which is also emphasized by literature [51]–[53].

III. THEORETICAL MODEL

In this section, we theoretically develop a profit maximizing model for the bank's optimal staffing policy under the seasonal demand variation. The first step is to abstract the bank's counter service system into a set of queueing models. Then, we define the functions of revenue and cost related to the seasonal staffing policy based on the queueing models. Finally, the profit maximizing model is constructed and simplified.

A. SERVICE SYSTEM IN THE BANK

Generally, we consider n branches in a commercial bank, denoted by $B_1, B_2, \cdots, B_n$ ($n \in \mathbb{N}^+$). To make the model as close to reality as possible, we use a seasonal M/M/c/N queueing system with balking and reneging to model the counter service system of each branch in the bank. For $i = 1, 2, \cdots, n$, the concrete definitions of the queueing system are as follows.

1) PEAK SEASON AND OFF-PEAK SEASON

The bank is faced with significant seasonal demand variation so that each year can be divided into a peak season and an off-peak season. For convenience, we define a logic variable t to distinguish peak season and off-peak season. $t = 0$ represents off-peak season while $t = 1$ represents peak season. And we use parameter T_t to denote the length of opening time of any branch in the bank in season t.

2) CUSTOMER ARRIVAL RATE

The difference between peak and off-peak season can be reflected by different numbers of customers who seek for service in the bank's branches per unit time. According to queueing theory [54], [55], we assume that customers arrive at B_i according to a Poisson process with an average arrival rate λ_{it}, which means there are averagely λ_{it} customers arriving at B_i per unit time in season t. Parameter λ_{it} mainly depends on the local market condition faced with the branch.

Exactly, λ_{it} measures the customer demands faced with B_i in season t, and $\lambda_{i1}/\lambda_{i0}$ measures the seasonal demand variation faced with B_i. An investigation conducted in China by us shows that λ_{i1} may reach three times as much as λ_{i0}.

3) SERVICE RATE

Each service takes some time. According to queueing theory [54], [55], we assume the service time as a random variable which follows the exponential distribution with a parameter μ_i, which is called service rate. It means that there are averagely μ_i customers receiving service in one service window in B_i per unit time. And $1/\mu_i$ is exactly the average service time. The value of μ_i is negatively affected by the average transaction size because a larger transaction tends to be more complex.

We do not separately consider the average service time in different seasons because μ_i is mainly affected by transaction complexity and service process length, while the influence of the customer flow is smaller.

4) SYSTEM CAPACITY

Obviously, only a limited number of customers can be accommodated in the business hall of each branch. We use parameter N_i to denote the maximum capacity of the area for customers in B_i (this capacity exactly refers to "N" in "M/M/c/N queueing system" [54], [55]). Let k denote the number of customers in the queueing system. Then, there must be $k \leq N_i$ in B_i. Newly arriving customers have no choice but to leave if $k = N_i$.

5) NUMBER OF SERVERS

To make the services match the customer demands in different seasons, the bank manager adopts a seasonal staffing policy as follows: arranging a group of permanent servers in each branch throughout the year, and sending additional staff from the head office to each branch in peak season. The core problem in this seasonal staffing policy is to determine how many permanent servers and additional servers for each branch.

For $i = 1, 2, \cdots, n$, let c_i denote the number of permanent servers in B_i, and δ_i denote the number of additional servers sent to B_i in the peak season. This denotation indicates there are c_i servers in B_i during the off-peak season, and $c_i + \delta_i$ servers in B_i during the peak season. Thus, the number of servers in B_i can be always expressed as $c_i + t\delta_i$ (this number of servers exactly refers to "c" in "M/M/c/N queueing system" [54], [55]).

For convenience, let

$$\boldsymbol{C} = (c_1, c_2, \cdots, c_n)$$
$$\boldsymbol{\Delta} = (\delta_1, \delta_2, \cdots, \delta_n)$$

Then $\boldsymbol{C}, \boldsymbol{\Delta}$ are decision variables for the bank manager. They should satisfy the following constraint conditions:

- **Constraint condition 1.** $(\boldsymbol{C}, \boldsymbol{\Delta}) \in \mathbb{N}^{2n}$ (the set of natural numbers) since they denote the numbers of persons.
- **Constraint condition 2.** $c_i + t\delta_i$ cannot be larger than the number of service windows in B_i, denoted by W_i. Thus, there must be $c_i + t\delta_i \leq W_i$ for $i = 1, 2, \cdots, n$ and $t = 0, 1$, which is equivalent to $c_i + \delta_i \leq W_i$ for $i = 1, 2, \cdots, n$.
- **Constraint condition 3.** To set up a mechanism of internal supervision, the bank must ensure that each branch has not less than 2 permanent servers. So, there must be $c_i \geq 2$ for $i = 1, 2, \cdots, n$.
- **Constraint condition 4.** Because of limited manpower, there is an upper bound of the number of staff available as additional servers for peak season. Letting H denote the upper bound, we have $\delta_1 + \delta_2 + \cdots + \delta_n \leq H$.

6) BALKING PROBABILITY

There must be some customers waiting in line when $c_i + t\delta_i \leq k \leq N_i$ (when the number of customers in B_i is not less than the number of servers). In this situation, a newly arriving customer may choose to balk and leave because he may not want to wait [27]–[29]. It is obvious that the probability of balking is positively correlated with k and negatively correlated with $c_i + t\delta_i$. According to the formula put forward by Abou-El-Ata and Hariri [29], we define this balking probability as follows:

$$b_{it}(c_i, \delta_i, k) = \begin{cases} 0, & 0 \leq k < c_i + t\delta_i \\ 1 - \dfrac{\beta\left(1 - \frac{k - c_i - t\delta_i + 1}{N_i}\right)}{k - c_i - t\delta_i + 2}, & c_i + t\delta_i \leq k < N_i \\ 1, & k = N_i \end{cases} \tag{1}$$

Here, β is a parameter to measure the newly arriving customer's willingness to join the line, which satisfies

$$0 \leq \beta \leq 2/\left(1 - 1/N_i\right)$$

7) RENEGING RATE

A customer waiting in line for service is also likely to give up waiting and choose to renege and leave [27]–[29]. Likewise, the rate of reneging is also positively correlated with k and negatively correlated with $c_i + t\delta_i$. According to the formula put forward by Abou-El-Ata and Hariri [29], we assume the reneging rate as follows:

$$\eta_{it}(c_i, \delta_i, k) = \begin{cases} 0, & 0 \leq k < c_i + t\delta_i \\ \alpha\,(k - c_i - t\delta_i), & c_i + t\delta_i \leq k \leq N_i \end{cases} \tag{2}$$

This rate means that the average number of reneging customers per unit time equals $\eta_{it}(c_i, \delta_i, k)$. Here, α is a positive parameter to measure the customer's willingness to renege while waiting.

8) PROBABILITY DISTRIBUTION OF NUMBER OF CUSTOMERS IN THE SERVICE SYSTEM

According to the formula put forward by Abou-El-Ata and Hariri [29], the probabilities of k customers being in B_i at any

$$p_{it}(c_i,\delta_i,k)=\begin{cases}\left\{\sum_{l=0}^{c_i+t\delta_i}\frac{\lambda_{it}^l}{\mu_i^l l!}+\frac{\lambda_{it}^{c_i+t\delta_i}}{\mu_i^{c_i+t\delta_i}(c_i+t\delta_i)!}\sum_{l=0}^{N_i-c_i-t\delta_i-1}\frac{C_{N_i-1}^{l+1}\beta^{l+1}\lambda_{it}^{l+1}}{(l+2)N_i^{l+1}\alpha^{l+1}\prod_{j=1}^{l+1}\left[\frac{\mu_i(c_i+t\delta_i)}{\alpha}+j\right]}\right\}^{-1}, & k=0\\ \frac{\lambda_{it}^k p_{it}(c_i,\delta_i,0)}{\mu_i^k k!}, & 0<k\le c_i+t\delta_i\\ \frac{\lambda_{it}^k\beta^{k-c_i-t\delta_i}C_{N_i-1}^{k-c_i-t\delta_i}p_{it}(c_i,\delta_i,0)}{\mu_i^{c_i+t\delta_i}\alpha^{k-c_i-t\delta_i}N_i^{k-c_i-t\delta_i}(k-c_i-t\delta_i+1)(c_i+t\delta_i)!\prod_{j=1}^{k-c_i-t\delta_i}\left[\frac{\mu_i(c_i+t\delta_i)}{\alpha}+j\right]}, & c_i+t\delta_i<k\le N_i\end{cases}\tag{3}$$

time point in season t for any $k=0,1,\cdots N_i$ can be figured out by (3), as shown at the top of this page.

B. REVENUE AND COST OF THE BANK

The bank manager seeks maximum profit when making decisions on $(\boldsymbol{C},\boldsymbol{\Delta})$. As profit equals revenue minus cost, then we construct the revenue function and cost function under the seasonal staffing policy. Because of the randomness of customer flow, the revenue and cost are both random variables. Thus, we consider the expectations of them so that the expected profit maximization model can be constructed. The variance is not taken into consideration because it will lead to an overly complex model.

Note that our seasonal staffing policy just covers counter service, so we do not consider the revenue and cost from online service channel and self-service channel.

1) EXPECTED REVENUE

Let R denote the total annual revenue from all transactions in $B_1,B_2,\cdots,B_n$. Then, we derive the expectation of R as a function of decision variables $\boldsymbol{C},\boldsymbol{\Delta}$.

For $i=1,2,\cdots,n$, let parameter R_i denote the average revenue from each transaction in branch B_i and in the whole year. These parameters are determined by factors like transaction size, business type and ability to bargain.

For $i=1,2,\cdots,n$ and $t=0,1$, let v_{it} denote the expected number of customers who accept services in B_i per unit time in season t. It is a random variable depending on the queueing system. According to the law of total probability, we have the following:

$$\mathrm{E}(v_{it})=\sum_{k=0}^{N_i}p_{it}(c_i,\delta_i,k)\,\mathrm{E}(v_{it}\,|k)\tag{4}$$

Here, $\mathrm{E}(v_{it}\,|k)$ is exactly equal to the expected number of customers arriving at B_i without balking or reneging when there are k customers in the system. Thus, we have

$$\mathrm{E}(v_{it}\,|k)=(1-b_{it}(c_i,\delta_i,k))\,\lambda_{it}-\eta_{it}(c_i,\delta_i,k)\tag{5}$$

Then, R can be expressed as a function of R_i, v_{it} and T_t ($i=1,2,\cdots,n$, $t=0,1$), as follow:

$$R=\sum_{t=0}^{1}\sum_{i=1}^{n}R_i v_{it}T_t\tag{6}$$

Here, v_{it} is a random variable while R_i and T_t are constant parameter. Thus, the expectation of R is

$$\mathrm{E}(R)=\sum_{t=0}^{1}\sum_{i=1}^{n}R_iT_t\mathrm{E}(v_{it})\tag{7}$$

By combining equations (1)-(5) and (7), we can express $\mathrm{E}(R)$ as a function of $\boldsymbol{C},\boldsymbol{\Delta}$ (there is no need to present this very big formula here).

2) EXPECTED COST

As discussed in Section II, what we mainly consider in the cost function under the seasonal staffing policy is the salary, which is composed of the basic salary and performance-related pay [56]. For any employee, the basic salary is constant and doesn't differ from one branch to another. Let S_1 denote the permanent server's average basic salary per unit time, and S_2 denote the additional server's average basic salary per unit time.

As for the performance-related pay, it should be in direct proportion to the employee's contribution to the bank. As mentioned above, we just need to consider how much revenue an employee brings to the bank through counter business. Thus, for $i=1,2,\cdots,n$ and $t=0,1$, the average contribution of a server per unit time in B_i in season t can be measured by $R_iv_{it}\big/(c_i+t\delta_i)$, i.e., total revenue from counter business in B_i per unit time divided by the number of servers. Then the average performance-related pay per unit time in B_i in season t can be expressed as $\gamma R_iv_{it}\big/(c_i+t\delta_i)$, where γ is a positive proportionality coefficient.

Therefore, the total salary per unit time in B_i in season t is as follow:

$$\begin{aligned}C_{it}&=\left[\left(S_1+\gamma\frac{R_iv_{it}}{c_i+t\delta_i}\right)c_i+\left(S_2+\gamma\frac{R_iv_{it}}{c_i+t\delta_i}\right)t\delta_i\right]T_t\\&=(S_1c_i+tS_2\delta_i+\gamma R_iv_{it})\,T_t\end{aligned}\tag{8}$$

Let C denote the total annual cost from servers' salaries in $B_1,B_2,\cdots,B_n$. We have

$$C=\sum_{t=0}^{1}\sum_{i=1}^{n}(S_1c_i+tS_2\delta_i+\gamma R_iv_{it})\,T_t\tag{9}$$

Here, v_{it} is a random variable while c_i and δ_i are non-random decision variables and S_1, S_2, γ, R_i and T_t are constant parameter. Thus, the expectation of C is

$$\mathrm{E}(C) = \sum\nolimits_{t=0}^{1} \sum\nolimits_{i=1}^{n} [S_1 c_i + tS_2\delta_i + \gamma R_i \mathrm{E}(v_{it})] T_t \quad (10)$$

Likewise, by combining equations (1)-(5) and (10), we can express $\mathrm{E}(C)$ as a function of $\boldsymbol{C}, \boldsymbol{\Delta}$ (there is no need to present this very big formula here).

C. PROFIT MAXIMIZING MODEL

The total annual expected profit equals $\mathrm{E}(R-C) = \mathrm{E}(R) - \mathrm{E}(C)$. Combining equations (4), (5), (7) and (10), we obtain

$$\begin{aligned}
&\mathrm{E}(R-C)\\
&= \sum_{t=0}^{1}\sum_{i=1}^{n} [(1-\gamma) R_i E(v_{it}) - S_1 c_i - tS_2\delta_i] T_t\\
&= \sum_{t=0}^{1}\sum_{i=1}^{n} \Bigg\{ (1-\gamma) R_i \sum_{k=0}^{N_i} p_{it}(c_i, \delta_i, k) [(1-b_{it}(c_i,\delta_i,k))\lambda_{it}\\
&\quad - \eta_{it}(c_i,\delta_i,k)] - S_1 c_i - tS_2\delta_i \Bigg\} T_t \quad (11)
\end{aligned}$$

It is apparent from equation (11) that $\mathrm{E}(R-C)$ can be expressed as a function of $\boldsymbol{C}, \boldsymbol{\Delta}$. For clarity, we set

$$\begin{aligned}
g_{it}(c_i,\delta_i) = &\Bigg\{ (1-\gamma) R_i \sum_{k=0}^{N_i} p_{it}(c_i,\delta_i,k) [(1-b_{it}(c_i,\delta_i,k))\lambda_{it}\\
&- \eta_{it}(c_i,\delta_i,k)] - S_1 c_i - tS_2\delta_i \Bigg\} T_t \quad (12)
\end{aligned}$$

and

$$f_0(\boldsymbol{C}, \boldsymbol{\Delta}) = \sum_{t=0}^{1}\sum_{i=1}^{n} g_{it}(c_i,\delta_i) \quad (13)$$

Thus, there is $f_0(\boldsymbol{C}, \boldsymbol{\Delta}) = \mathrm{E}(R-C)$.

Since the commercial bank is a for-profit institution, the bank manager seeks to make the seasonal staffing policy as profitable as possible. Thus, $f_0(\boldsymbol{C}, \boldsymbol{\Delta})$ is exactly the objective function of the bank manager's decision problem when he makes the seasonal staffing policy.

The constraint conditions of the decision variables have been listed in subsection III-A-e. Then, we can write down the bank manager's optimization model as follows:

$$\begin{aligned}
\mathrm{M}_0: \quad &\max f_0(\boldsymbol{C}, \boldsymbol{\Delta}) = \sum_{t=0}^{1}\sum_{i=1}^{n} g_{it}(c_i,\delta_i)\\
&\text{subject to } c_i, \delta_i \in \mathbb{N}, \quad i = 1, 2, \cdots, n\\
&\qquad c_i \geq 2, \quad i = 1, 2, \cdots, n\\
&\qquad c_i + \delta_i \leq W_i, \quad i = 1, 2, \cdots, n\\
&\qquad \delta_1 + \delta_2 + \cdots + \delta_n \leq H
\end{aligned}$$

D. LINEARIZATION OF THE MODEL

The formula of $g_{it}(c_i,\delta_i)$ (equation (12)) contains $b_{it}(c_i,\delta_i,k)$, $\eta_{it}(c_i,\delta_i,k)$ and $p_{it}(c_i,\delta_i,k)$ that are nonlinear functions of c_i and δ_i. This means M_0 is a nonlinear integer programming problem that is difficult to solve directly. Thus, we need to linearize model M_0 so as to simplify it.

1) LINEARIZATION OF OBJECTIVE FUNCTION

Inspired by the theory of linearization technique for solving discrete problem [44], [45], we define the auxiliary variables to linearize the objective function of M_0. For $i = 1, 2, \cdots, n$, $r = 2, 3, \cdots, W_i$ and $s = 0, 1, \cdots, W_i$, we define

$$z_{irs} = \begin{cases} 1, & (r = c_i) \wedge (s = \delta_i)\\ 0, & \text{else} \end{cases} \quad (14)$$

Then we can use this group of auxiliary 0-1 variables to transform equation (13), as follow:

$$\sum_{t=0}^{1}\sum_{i=1}^{n} g_{it}(c_i,\delta_i) = \sum_{t=0}^{1}\sum_{i=1}^{n}\sum_{r=2}^{W_i}\sum_{s=0}^{W_i} z_{irs} g_{it}(r,s) \overset{def}{=} f_1(\mathbf{Z}) \quad (15)$$

where

$$\begin{aligned}
\mathbf{Z} &= (\mathbf{Z}_1, \mathbf{Z}_2, \cdots, \mathbf{Z}_n)\\
\mathbf{Z}_i &= (\mathbf{Z}_{i2}, \mathbf{Z}_{i3}, \cdots, \mathbf{Z}_{iW_i})\\
\mathbf{Z}_{ir} &= (z_{ir0}, z_{ir1}, \cdots, z_{irW_i})
\end{aligned}$$

To complete the linearization, we need to find an equivalent definition of z_{irs} that doesn't rely on c_i and δ_i. Theorem 1 gives such an equivalent definition:

Theorem 1: There is one-to-one mapping between $(\boldsymbol{C}, \boldsymbol{\Delta})$ and $\mathbf{Z}$, if $\mathbf{Z}$ satisfies the following conditions, for $i = 1, 2, \cdots, n$:

$$\begin{cases} z_{irs} \in \{0, 1\}, \quad r = 2, 3, \cdots, W_i, \ s = 0, 1, \cdots, W_i\\ \sum\nolimits_{r=2}^{W_i} \sum\nolimits_{s=0}^{W_i} z_{irs} = 1 \end{cases} \quad (16)$$

The one-to-one mapping is formulated as follows:

$$c_i = \sum\nolimits_{r=2}^{W_i} \sum\nolimits_{s=0}^{W_i} r z_{irs} \quad (17)$$

$$\delta_i = \sum\nolimits_{r=2}^{W_i} \sum\nolimits_{s=0}^{W_i} s z_{irs} \quad (18)$$

Proof: For $i = 1, 2, \cdots, n$, it can be easily inferred from conditions (16) that: (a) there is unique $r' \in \{2, 3, \cdots, W_i\}$ and $s' \in \{0, 1, \cdots, W_i\}$ satisfying $z_{ir's'} = 1$; (b) for all $r \neq r'$ and $s \neq s'$, there is $z_{irs} = 0$. Letting $c_i = r'$ and $\delta_i = s'$, we obtain the one-to-one mapping between (c_i, δ_i) and $\mathbf{Z}_i$ that is formulated as equations (17) and (18). □

According to Theorem 1, conditions (16) exactly provide an equivalent definition of z_{irs} that doesn't rely on c_i and δ_i. Thus, conditions (16) and equation (15) give a complete linearization of the objective function of M_0.

2) CONSTRUCTION OF EQUIVALENT LINEAR MODEL

Substituting equations (17) and (18) into the constraint conditions of M_0, we transform these constraint conditions into the following

$$\begin{cases} \sum_{r=2}^{W_i}\sum_{s=0}^{W_i} rz_{irs} \geq 2, \quad i=1,2,\cdots,n \\ \sum_{r=2}^{W_i}\sum_{s=0}^{W_i}(r+s)\,z_{irs} \leq W_i, \quad i=1,2,\cdots,n \\ \sum_{i=1}^{n}\sum_{r=2}^{W_i}\sum_{s=0}^{W_i} sz_{irs} \leq H \end{cases} \tag{19}$$

Therefore, the optimization model M_0 can be transformed into a completely equivalent model M_1, which is a 0-1 type linear integer programming model as follows (see equation (12) for the expression of function g_{it}):

$$M_1: \ \max f_1(\mathbf{Z}) = \sum_{t=0}^{1}\sum_{i=1}^{n}\sum_{r=2}^{W_i}\sum_{s=0}^{W_i} z_{irs} g_{it}(r,s)$$

$$\text{subject to } z_{irs} \in \{0,1\}, \quad i=1,2,\cdots,n, \quad r=2,3,\cdots, \quad W_i, s=0,1,\cdots,W_i$$

$$\sum_{r=2}^{W_i}\sum_{s=0}^{W_i} z_{irs} = 1, \quad i=1,2,\cdots,n$$

$$\sum_{r=2}^{W_i}\sum_{s=0}^{W_i} rz_{irs} \geq 2, \quad i=1,2,\cdots,n$$

$$\sum_{r=2}^{W_i}\sum_{s=0}^{W_i}(r+s)\,z_{irs} \leq W_i, \quad i=1,2,\cdots,n$$

$$\sum_{i=1}^{n}\sum_{r=2}^{W_i}\sum_{s=0}^{W_i} sz_{irs} \leq H$$

Due to the equivalence relation between M_0 and M_1, we can obtain the following theorem:

Theorem 2: Let $(\boldsymbol{C}^*, \boldsymbol{\Delta}^*)$ denote the solution vector of M_0, and $\mathbf{Z}^*$ denote the solution vector of M_1. There must be the following:

$$f_0(\boldsymbol{C}^*, \boldsymbol{\Delta}^*) = f_1(\mathbf{Z}^*) \tag{20}$$

And for $i=1,2,\cdots,n$, there are

$$c_i^* = \sum_{r=2}^{W_i}\sum_{s=0}^{W_i} rz_{irs}^* \tag{21}$$

$$\delta_i^* = \sum_{r=2}^{W_i}\sum_{s=0}^{W_i} sz_{irs}^* \tag{22}$$

where c_i^*, δ_i^*, z_{irs}^* are respectively the element of $\boldsymbol{C}^*$, $\boldsymbol{\Delta}^*$, $\mathbf{Z}^*$.

Thus, by solving M_1 and substituting its solution into equations (21) and (22), the optimal values of c_i and δ_i ($i = 1,2,\cdots,n$) can be found. Since f_1, the objective function of M_1, is a linear function of decision vectors $\mathbf{Z}$, M_1 can be solved by using the implicit enumeration method, a special branch and bound method suited for solving zero-one type linear integer programming problems. Please refer to references [46]–[48] for a detailed introduction of the method. It seems impossible to obtain an analytical solution due to the complexity of M_1 despite the linearization. However, we are still able to quickly get a numerical solution with the help of a computer program. In fact, the numerical solution is absolutely precise since the solution is a set of integers. An available tool for solving M_1 is a mathematical programming solver named "Gurobi", which supports various programming languages, including C, MATLAB, C++, Java, Python, .Net, etc. Or you can use the mixed-integer linear programming solver named "intlinprog" in MATLAB.

E. PRACTICAL APPLICATION

In this subsection, we explain how to apply model M_1 to make the optimal seasonal staffing policy. The method includes four parts: data collection, parameter estimation, programming, and implementation. There are 13 specific steps as follows:

Step 1. Distinguish between the peak season and off-peak season in the year based on transaction data collected from branches so that the parameter T_t ($t = 0, 1$) can be determined.

Step 2. Estimate λ_{it} by the data of the average customers arrival per unit time, μ_i by the data of average service times per unit time, and R_i by the data of average return per transaction ($i = 1,2,\cdots,n, t = 0,1$).

Step 3. Based on actual observations, set α and β equal to values in our recommended ranges: $\alpha \in [0.05, 0.25]$, $\beta \in [0.5, 1]$. A survey is helpful to estimate α and β more precisely, but it is so costly that we do not recommend using it.

Step 4. Determine the values of parameters W_i, H, N_i, S_1, S_2, γ ($i = 1,2,\cdots,n$) by simple observation and counting.

Step 5. Create a computer program in C++, MATLAB or Python.

Step 6. Define the auxiliary variables z_{irs} ($i = 1,2,\cdots,n$, $r = 2,3,\cdots,W_i$, $s = 0,1,\cdots,W_i$) in the program according to equation (16).

Step 7. Define c_i and δ_i ($i = 1,2,\cdots,n$) as functions of z_{irs} ($i = 1,2,\cdots,n$, $r = 2,3,\cdots,W_i$, $s = 0,1,\cdots,W_i$) according to equations (17) and (18).

Step 8. Define the parameters T_t, λ_{it}, μ_i, R_i, α, β, W_i, H, N_i, S_1, S_2, γ ($i = 1,2,\cdots,n$, $t = 0,1$) in the program and input their values.

Step 9. Define the functions $p_{it}(r,s,k)$, $b_{it}(r,s,k)$ and $\eta_{it}(r,s,k)$ in the program.

Step 10. Write down the constraint conditions and the objective function of M_1 in the program.

Step 11. Call the mathematical programming solver "Gurobi" or "intlinprog" to calculate the optimal values of z_{irs} ($i = 1,2,\cdots,n$, $r = 2,3,\cdots,W_i$, $s = 0,1,\cdots,W_i$), which leads to the optimal values of c_i and δ_i ($i = 1,2,\cdots,n$).

Step 12. Design the seasonal staffing policy based on the program output.

Step 13. Implement the seasonal staffing plan.

It is neither time-consuming nor costly to execute the above procedures, so it is convenient for the commercial bank to implement our method every year to meet the change of market condition. Note that our method is also applicable for other service institutions with branches that are faced with seasonal demand variations, but some adjustments might be needed to fit the actual situation.

IV. NUMERICAL EXPERIMENTS

In this section, we not only explore the effects of parameters on the optimal seasonal staffing policy derived from model M_1, but also compare our seasonal staffing method based on M_1 with the traditional methods of dealing with seasonal demand variation, including the rough seasonal staffing method (seasonal staffing based on rough estimations) and the unseasonal staffing method.

A. PREPARATIONS

Before the numerical experiments, we provide the concrete definitions of the two traditional staffing methods of dealing with seasonal demand variation.

1) ROUGH SEASONAL STAFFING METHOD

Under this method, the manager roughly estimates the required numbers of permanent servers and additional servers in each branch based on the average customer arrival rate in either season and the average service rate. According to queueing theory [54], [55], a simple way of estimating the required number of servers is to use the following:

$$c = \left\lceil \frac{\lambda}{\mu} \right\rceil$$

where $\lceil \cdot \rceil$ returns the nearest integer larger than or equal to the element.

At the same time, the manager must ensure c_i and δ_i to satisfy the constraint conditions given in Section III-A-e. Thus, we construct the following seasonal staffing method based on rough estimations (for $1, 2, \cdots, n$):

$$\begin{cases} c_i^R = \min\left\{\max\left\{\left\lceil \frac{\lambda_{i0}}{\mu_i} \right\rceil, 2\right\}, W_i\right\} \\ \delta_i^R = \max\left\{\min\left\{\left\lceil \frac{\lambda_{i1}}{\mu_i} \right\rceil - c_i^R, W_i - c_i^R, \left\lfloor \frac{\left(\left\lceil \frac{\lambda_{i1}}{\mu_i} \right\rceil - c_i^R\right)H}{\sum_{j=1}^{n}\left(\left\lceil \frac{\lambda_{j1}}{\mu_j} \right\rceil - c_i^R\right)} \right\rfloor\right\}, 0\right\} \end{cases} \tag{23}$$

where $\lfloor \cdot \rfloor$ returns the nearest integer less than or equal to the element.

Then it is not hard to get the following theorem.

Theorem 3: For $1, 2, \cdots, n$, c_i^R and δ_i^R calculated by (23) must satisfy the following constraint conditions:

$$c_i^R \geq 2$$
$$c_i^R + \delta_i^R \leq W_i$$
$$\delta_1^R + \delta_2^R + \cdots + \delta_n^R \leq H$$

Let $\left(\mathbf{C^R}, \mathbf{\Delta^R}\right) = \left(c_1^R, c_2^R, \cdots, c_n^R, \delta_1^R, \delta_2^R, \cdots, \delta_n^R\right)$. Then $f_0\left(\mathbf{C^R}, \mathbf{\Delta^R}\right)$ exactly measures the performance of the rough seasonal staffing method.

2) UNSEASONAL STAFFING METHOD

The unseasonal staffing means allocating a fixed number of servers in each branch in the whole year, regardless of the seasonal demand variation. The manager just estimates the required number of permanent servers in each branch based on the average customer arrival rate in the whole year and the average service rate. Similarly, the unseasonal staffing method can be constructed as follow (for $1, 2, \cdots, n$):

$$c_i^U = \min\left\{\max\left\{\left\lceil \frac{\lambda_{i0}T_0 + \lambda_{i0}T_1}{\mu_i\left(T_0 + T_1\right)} \right\rceil, 2\right\}, W_i\right\} \tag{24}$$

And we have the following theorem.

Theorem 4: For $1, 2, \cdots, n$, c_i^U calculated by (24) must satisfy the following constraint conditions:

$$c_i^U \geq 2$$
$$c_i^U \leq W_i$$

Let $\mathbf{C^U} = \left(c_1^U, c_2^U, \cdots, c_n^U\right)$. Then $f_0\left(\mathbf{C^U}, \mathbf{0}\right)$ exactly measures the performance of the unseasonal staffing method.

Now we have three staffing methods under seasonal demand variation, as Fig.2 shows:

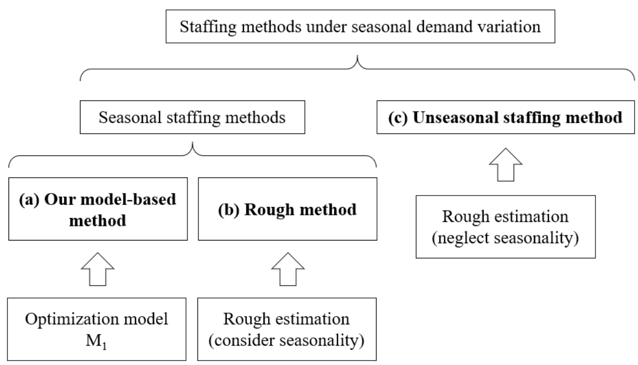


FIGURE 2. Three staffing methods under seasonal demand variation.

The latter two are used by most of commercial banks. Then we conduct a set of numerical experiments to show the superiority of our method, and further explore the impacts of factors on the performances of the methods.

B. EXPERIMENT DESCRIPTIONS

We consider 5 branches in the bank ($n = 5$). The first 4 branches B_1, B_2, B_3, B_4 constitute the reference group, while the last one B_5 serves as the test object. The 4 branches in the reference group correspond to 4 types of branches under the framework of our model, as showed by Table 1.

TABLE 1. A classification of branches in commercial banks.

Criterion	Type 1	Type 2	Type 3	Type 4
Demand quantity	High	High	Low	Low
Transaction size	High	Low	High	Low

Here, demand quantity is positively correlated with parameter λ_{it} and transaction size is positively correlated with parameter R_i.

For $i = 1, 2, \cdots, 4$ and $t = 0, 1$, parameters λ_{it}, R_i, μ_i, W_i and N_i are reference parameters that are set to be

TABLE 2. Parameters setting of the reference group.

Parameter	$i=1$	$i=2$	$i=3$	$i=4$
$\lambda_{it}, t=0$	48	48	24	24
$\lambda_{it}, t=1$	96	96	48	48
R_i	2	1	2	1
μ_i	6	12	6	12
W_i	8	8	4	4
N_i	40	40	20	20

TABLE 3. Data of customer arrival.

Month	Average Number of Customers per Hour	
	Bank 1	Bank 2
January	91.2	52.0
February	102.7	54.9
March	87.5	47.7
April	50.0	25.1
May	44.9	23.9
June	47.2	26.5
July	55.7	27.1
August	57.1	30.1
September	45.8	24.6
October	48.8	25.3
November	51.4	27.8
December	90.3	43.4

always constants. See Table 2 for their values in detail (the unit of R_i is 100 US dollars and the unit of time is hours).

Here the parameter values are set based on surveys in two local commercial banks in China. Table 3 presents the data collected in the surveys, including the data of customer arrival in two banks during year 2017. The data is not enough to support an empirical study so that it just serves as a reference for parameter setting.

The test parameters should have involved λ_{5t}, μ_5, R_5, W_5, N_5, T_t, α, β, H, S_1, S_2, γ $(t=0,1)$. However, some of them can be eliminated or replaced, as follows:

1) T_0+T_1 is obviously constant so that we can eliminate T_0 and just take T_1 into the sensitivity test ($T_0+T_1=1440$ hours).
2) For $i=1,2,\cdots,5$, R_i tends to be positively correlated with the average transaction size, while μ_i is negatively affected by the average transaction size. Therefore, μ_i tends to be negatively correlated with R_i. We assume $\mu_i R_i = 1200$ so that μ_5 can be eliminated from the test parameters.
3) The ratio $\lambda_{51}/\lambda_{50}$ is likely to be an important factor of the staffing policy. Letting $\lambda_{50}=\lambda_5$ and $\lambda_{51}=\sigma_5\lambda_5$, we treat λ_5 and σ_5 as the parameters to be tested. This means λ_{51} and λ_{50} are changing synchronously when we change λ_5, yet λ_{50} is fixed when we change σ_5.
4) For a similar reason, we let $S=S_1$ and $\zeta=S_2/S_1$, and treat S and ζ as the parameters to be tested.

Through the eliminations and replacements above, we obtain a new group of test parameters:

$$\Theta = \{\lambda_5, \sigma_5, R_5, W_5, N_5, T_1, \alpha, \beta, H, S, \zeta, \gamma\}$$

In the next subsection we will explore how $(\boldsymbol{C}^*, \boldsymbol{\Delta}^*)$, $(\boldsymbol{C}^{\mathbf{R}}, \boldsymbol{\Delta}^{\mathbf{R}})$, $\boldsymbol{C}^{\mathbf{U}}$, f_0^*, f_0^R, and f_0^U change with each parameter

TABLE 4. Ranges and initial values of test parameters.

Parameter	Initial value	Range
λ_5	36	[0,80]
σ_5	2	[1,3]
R_5	1.5	[0.5,2.5]
W_5	6	[2,11]
N_5	30	[10,50]
T_1	480	[0,1440]
α	0.15	(0,1)
β	1	[0,2]
H	15	[5,25]
S	5	[1,9]
ζ	1	[0.5,1.5]
γ	0.2	[0,0.4]

in Θ. When we change the value of one parameter, the others are set to the initial values. The ranges and initial values of the test parameters are given by Table 4. All the numerical experiments are conducted under the help of MATLAB.

C. RESULTS AND ANALYSES

For convenience, we let $f_0^* = f_0(\boldsymbol{C}^*, \boldsymbol{\Delta}^*)$, $f_0^R = f_0(\boldsymbol{C}^{\mathbf{R}}, \boldsymbol{\Delta}^{\mathbf{R}})$, and $f_0^U = f_0(\boldsymbol{C}^{\mathbf{U}}, \mathbf{0})$. Then we use curve charts to present the effects that some key parameters in Θ have on f_0^*, f_0^R, and f_0^U, and to compare f_0^*, f_0^R, and f_0^U under different parameter values. And some parameters significantly affect (c_5^*, δ_5^*), (c_5^R, δ_5^R), and c_5^U. We show these impacts by bar charts. Due to the limited space, we do not show the changes of (c_i^*, δ_i^*), (c_i^R, δ_i^R), and c_i^U for 1, 2, 3, 4, but some necessary descriptions are still provided.

1) EFFECTS OF DEMAND QUANTITY (MEASURED BY λ_5)

The trends of f_0^*, f_0^R, and f_0^U with increasing λ_5 are shown by Fig.3.

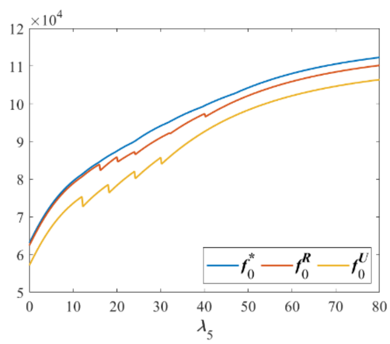


FIGURE 3. Effect of λ_5 on f_0^*, f_0^R, and f_0^U.

It can be seen that: (1) there is always $f_0^* > f_0^R > f_0^U$ for any $\lambda_5 \in [0, 80]$, but $f_0^* - f_0^R$ is relatively slight when λ_5 is at a low level; (2) as λ_5 increases, f_0^* grows with a marginal diminishing effect, while f_0^R and f_0^U increase with marginal diminishing effects except at a few discontinuous points; (3) at the discontinuity points on the curves of f_0^R and f_0^U, the curves always jump down.

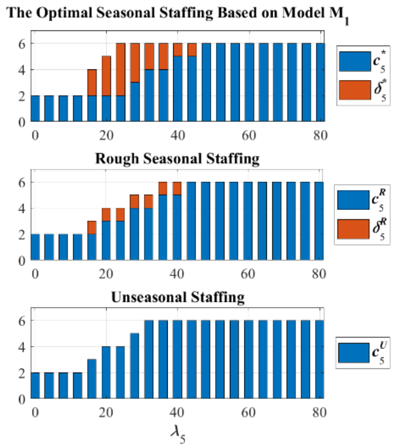


FIGURE 4. Effect of λ_5 on $\left(c_i^*, \delta_i^*\right)$, $\left(c_i^R, \delta_i^R\right)$, and c_i^U.

As long as the customer demand is not too small, our seasonal staffing method based on model M_1 has significant advantage over the rough seasonal staffing method and the unseasonal one. Although $\left(c_5^R, \delta_5^R\right)$ is equal to $\left(c_5^*, \delta_5^*\right)$ when λ_5 is at a high level (see Fig.4), f_0^* is still larger than f_0^R because of a difference between δ_2^* and δ_2^R. This indicates that our method is capable of achieving optimal human resource allocation in all branches, while the rough seasonal staffing method is difficult to achieve global optimization though it may reach local optimization in some branches.

The discontinuity points on the curves of f_0^R and f_0^U are caused by the change of staffing — the integer changes by jumping. The curves always jump down since that an upward jump of $\left(c_5^R, \delta_5^R\right)$ or c_5^U is likely to bring an upward jump of salaries and a smaller upward jump of business incomes. By contrast, our method is based on an optimization model so that it can find those critical points to avoid jumps through optimization. This indicates that our method can avoid adverse fluctuations of profit when customer demand changes.

The marginal diminishing effects on the curves of f_0^*, f_0^R, and f_0^U respect to λ_5 are attributed to two reasons: on the one hand, as the customer flow increases, the rising rate of balking and reneging results in a diminishing growth of transactions; on the other hand, since the number of servers must not exceed the number of service windows ($c_5 + \delta_5 \leq W_5$), it is impossible to arrange more servers to satisfy the increasing customer flow in B_5 when all the service windows are occupied.

Actually, if at a very low level, parameters H or N_5 may also strengthen the marginal diminishing effects on the curves of f_0^*, f_0^R, and f_0^U respect to λ_5, because small H means a shortage of staff members in the head office available for transferring in peak season, and low N_5 means the service system will be under a full load even though λ_5 is not too large.

Then let's look at the changes of the three staffing policies with increasing λ_5, as Fig.4 shows.

We can see that: (1) the number of servers is positively correlated with λ_5 in a certain range under any of the three staffing methods; (2) the number of servers stop increasing with λ_5 when reaching the number of service windows; (3) the number of servers under the seasonal staffing method based on model M_1 reaches the upper bound earlier than that under the rough seasonal staffing method and the unseasonal one; (4) under any of the two seasonal staffing methods, there is a space to arrange additional servers for peak season when λ_5 is in a certain range, but the space is eliminated when λ_5 is beyond the range; (5) the seasonal staffing method based on model M_1 tends to allocate more additional servers than the rough seasonal staffing method when there is space for additional servers.

Compared with the rough seasonal staffing method, our method tends to arrange more additional servers for peak season if there is enough space for them. This indicates a higher flexibility of our method in human resource allocation. It's another reason for the better performance of our seasonal staffing method.

The evidence of the limitation of the number of servers (W_5) is apparent in Fig.4. Moreover, when we increase W_5 from 6 to 12 (raise the upper bound of servers), the space for servers is enlarged so that the marginal diminishing effects on the growth trends of f_0^*, f_0^R, and f_0^U with respect to λ_5 are weakened. Fig.5 and Fig.6 show the changes. This result indicates a very important role of the number of service windows in the issue of staffing.

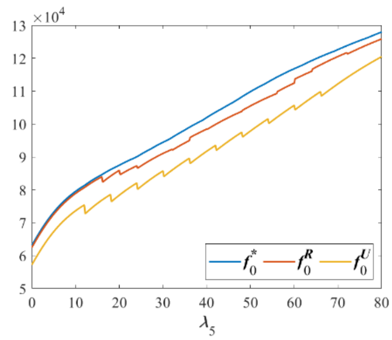


FIGURE 5. Effect of λ_5 on f_0^*, f_0^R, and f_0^U ($W_5 = 12$).

Besides, note that effects of parameter R_5 are similar with that of λ_5 which are analyzed above, so we do not provide repetitive descriptions and analyses on the effects of R_5.

2) EFFECTS OF SEASONAL DEMAND VARIATION (MEASURED BY σ_5)

Fig.7 shows the influences of σ_5 on f_0^*, f_0^R, and f_0^U.

It can be seen that: (1) there is always $f_0^* > f_0^R > f_0^U$ for any $\sigma_5 \in [1, 5]$; (2) as σ_5 increases, f_0^* presents a growth trend with a marginal diminishing effect, while f_0^R and f_0^U increase with marginal diminishing effects except at the discontinuous points.

The reasons and implications of the trends in Fig.7 are similar with that of the trends in Fig.3, so we do not give

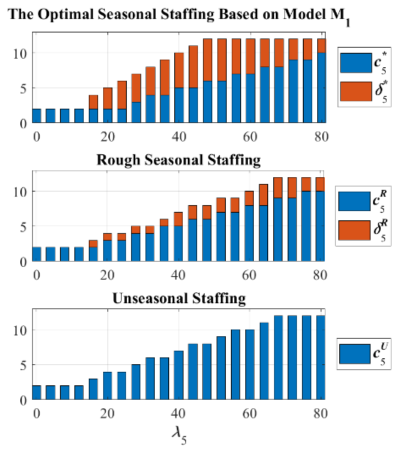


FIGURE 6. Effect of λ_5 on $\left(c_i^*, \delta_i^*\right)$, $\left(c_i^R, \delta_i^R\right)$, and c_i^U ($W_5 = 12$).

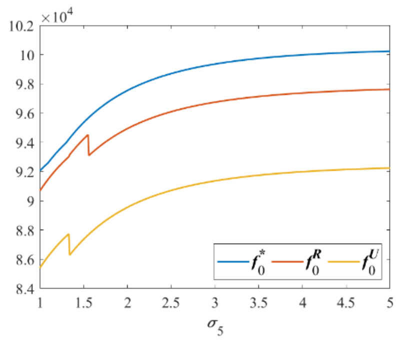


FIGURE 7. Effect of σ_5 on f_0^*, f_0^R, and f_0^U.

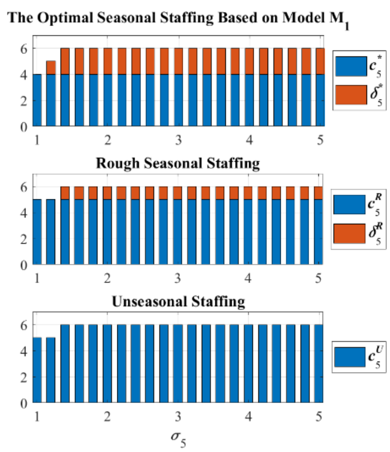


FIGURE 8. Effect of σ_5 on $\left(c_i^*, \delta_i^*\right)$, $\left(c_i^R, \delta_i^R\right)$, and c_i^U.

a repetition. The main difference is that λ_5 concerns the customer demands in the whole year while σ_5 just concerns the customer demands in peak season, so σ_5 has slighter impacts on the three staffing policies (see Fig.8). This is the reason why the numbers of discontinuous points on the curves of f_0^R and f_0^U with respect to σ_5 are much fewer than that on the curves of f_0^R and f_0^U with respect to λ_5.

Note that, when at a low level, parameters H and N_5 may also strengthen the marginal diminishing effects on the curves of f_0^*, f_0^R, and f_0^U respect to σ_5. The explanations for this are similar with that provided in the analyses on the effects of λ_5 (in the previous page).

The changes of the three staffing policies with increasing σ_5 are showed by Fig.8.

We can see that: (1) all the three staffing policies just change when σ_5 is at a low level; (2) under the two seasonal staffing methods, the changes of staffing are exactly the changes of the numbers of additional servers for peak season, while the numbers of permanent servers remain constant; (3) our seasonal staffing method based on model M_1 tends to arrange more additional servers than the rough seasonal staffing method does, which indicates our method has higher flexibility of human resource allocation.

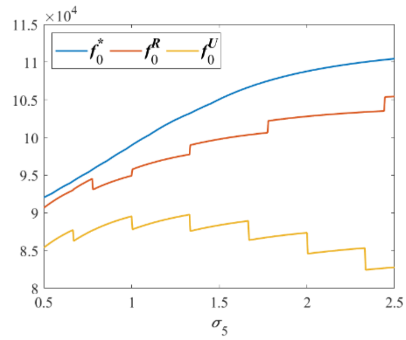


FIGURE 9. Effect of σ_5 on f_0^*, f_0^R, and f_0^U ($W_5 = 12$).

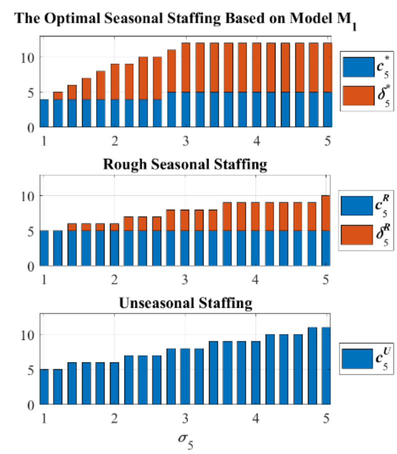


FIGURE 10. Effect of σ_5 on $\left(c_i^*, \delta_i^*\right)$, $\left(c_i^R, \delta_i^R\right)$, and c_i^U ($W_5 = 12$).

When increasing W_5 from 6 to 12, we obtain the results showed by Fig.9 and Fig.10. There are weaker diminishing effects on the curves of f_0^* and f_0^R in Fig.9. And Fig.10 shows a significantly enlarged space for staffing.

It can be seen from Fig.9 that the advantage of our method (measured by $f_0^* - f_0^R$ and $f_0^* - f_0^U$) is positively correlated with σ_5 approximately under a large W_5, which indicates that our method can reach high efficiency of human resource allocation when there is sufficient space for staffing.

Surprisingly, the curve of f_0^U in Fig.9 presents a downward trend in the second half of the curve. This result indicates a serious deficiency in effectiveness of human resource allocation of the unseasonal staffing method in the situation of high seasonal demand variation.

3) EFFECTS OF THE NUMBER OF SERVICE WINDOWS (MEASURED BY W_5)

The above results indicate an important role of W_5 in our problem. Now let's take a close look at the effects of W_5. Fig.11 shows its effects on f_0^*, f_0^R, and f_0^U.

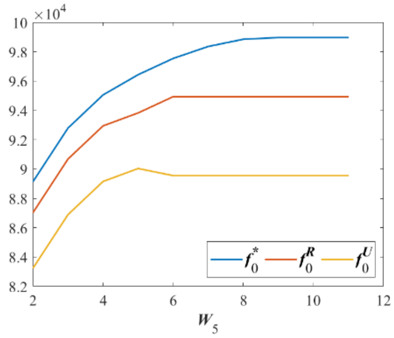


FIGURE 11. Effect of W_5 on f_0^*, f_0^R, and f_0^U.

From Fig.11 we see that: (1) there is always $f_0^* > f_0^R > f_0^U$ for any $W_5 \in \{2, 3, \cdots, 11\}$; (2) the curves of f_0^* and f_0^R first increase with significant marginal diminishing effects and finally level out; (3) the curve of f_0^U first increase with significant marginal diminishing effects, then drop slightly and finally level out; (4) $f_0^* - f_0^R$ and $f_0^* - f_0^U$ increase with W_5 before f_0^* stop increasing.

The reason for the diminishing and eventually disappeared growth trends on the curves of f_0^*, f_0^R, and f_0^U with respect to W_5 is that there is no need to add more service windows for more servers when customer demands are satisfied well. The upper bounds of the curves of f_0^* and f_0^R with respect to W_5 may be reduced if H or N_5 is very small — if there is no staff members available for transferring and the number of permanent servers is enough to meet the customer demand in off-peak seasonal, it is naturally useless to add service windows for more additional servers for peak season; and if the capacity of the service is too small, the system is easy to reach full load under which adding service windows is obviously meaningless.

Then let's look at the changes of $\left(c_5^*, \delta_5^*\right)$, $\left(c_5^R, \delta_5^R\right)$, and c_5^U with increasing W_5, showed by Fig.12. Compared to the rough seasonal staffing method and the unseasonal one, our method tends to arrange fewer permanent servers and more additional servers for peak season, indicating the higher flexibility of human resource allocation of our method.

Returning to the theoretical model, we discuss the above phenomenon from a more general view of point. Determined by the principle of service resources matching customer demands, the optimal solution of M_0 always exists within a certain range even though W_i is set to a very large value ($i = 1, 2, \cdots, n$). Thus, the following corollary can be obtained:

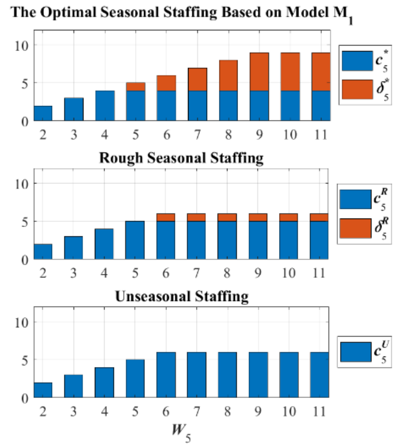


FIGURE 12. Effect of W_5 on $\left(c_i^*, \delta_i^*\right)$, $\left(c_i^R, \delta_i^R\right)$, and c_i^U.

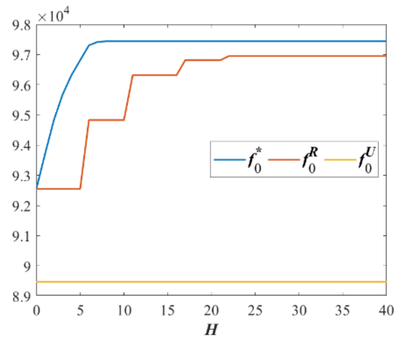


FIGURE 13. Effect of H on f_0^*, f_0^R, and f_0^U.

Corollary 1: Let $\left(\boldsymbol{C}^{**}, \boldsymbol{\Delta}^{**}\right) = \left(c_1^{**}, c_2^{**}, \cdots, c_n^{**}, \delta_1^{**}, \delta_2^{**}, \cdots, \delta_n^{**}\right)$ denote the optimal solution of M_0 given $W_i = \infty$ for $i = 1, 2, \cdots, n$. Given other conditions being equal, the solutions of M_0 must satisfy that under the condition of $W_i \geq c_i^{**} + \delta_i^{**}$ (for $i = 1, 2, \cdots, n$), there is

$$\left(\boldsymbol{C}^{**}, \boldsymbol{\Delta}^{**}\right) = \left(\boldsymbol{C}^{*}, \boldsymbol{\Delta}^{*}\right)$$

This corollary implies the existence of the optimal W_i under our seasonal staffing method (denoted by W_i^*). It is not necessarily equal to $c_i^{**} + \delta_i^{**}$ since we have not taken the cost of installing service windows into consideration. However, it is easy to infer that W_i^* should be less than $c_i^{**} + \delta_i^{**}$ under the seasonal policy if we add this cost to the models. The issue of determining the number of service windows actually belongs to another subfield — physical resource allocation. It is also worth deeper research in the future.

4) EFFECT OF THE NUMBER OF STAFF MEMBERS AVAILABLE FOR TRANSFERRING IN PEAK SEASON (MEASURED BY H)

The trends of f_0^*, f_0^R, and f_0^U with increasing H are shown by Fig.13, in which we can see that: (1) there is always $f_0^* > f_0^R > f_0^U$ for any $H \in \{0, 1, \cdots, 25\}$; (2) the curve of f_0^* first increases with a marginal diminishing effect and then level out; (3) the curve of f_0^R jumps up on and off with diminishing jumping height (this can also be regarded as a marginal diminishing effect on the growth trend of f_0^R), and levels out after the fourth jump; (4) f_0^U is not affected by H at all.

Similar with W_5, parameter H has gradually diminish positive effects on f_0^* and f_0^R in certain ranges, and the effects disappear when H is beyond the ranges. This phenomenon is caused by the fact that there is no need to arrange more servers when customer demands are satisfied well. The upper bounds of the curves of f_0^* and f_0^R with respect to H may be reduced if each W_i or N_i $(i = 1, 2, \cdots, n)$ is very small — if there is no space for more servers, it is naturally useless to provide more staff members for transferring in peak season; and if the service system too small, it is easy to reach full load under which adding more servers cannot provide any help.

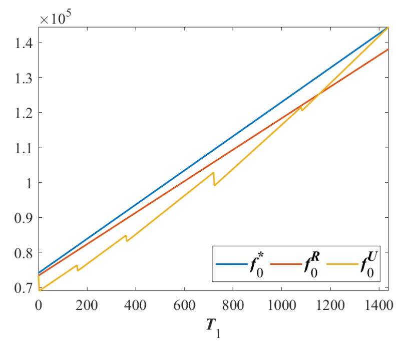


FIGURE 14. Effect of T_1 on f_0^*, f_0^R, and f_0^U.

5) EFFECT OF THE LENGTH OF THE PEAK SEASON (MEASURED BY T_1)

Fig.14 present the effects of T_1 on f_0^*, f_0^R, and f_0^U. (1) There are always $f_0^* > f_0^R$ and $f_0^* > f_0^U$ for any $T_1 \in (0, 1440)$, but there is $f_0^* = f_0^U$ when $T_1 = 0$ or 1440. (2) The curve of f_0^U with respect to T_1 increases discontinuously with marginal increasing effects. (3) f_0^U increases faster than f_0^R except at the discontinuity points, and there is $f_0^U > f_0^R$ when T_1 is at a high level. (4) There are directly proportional relationships between f_0^* and T_1 and between f_0^R and T_1. (5) The slope of the linear curve of f_0^* is larger than that of f_0^R.

The linearity of f_0^* and f_0^R with respect to T_1 is attributed to the linearity of f_0 with respect to T_1 in addition to no change of staffing policies under the two seasonal staffing methods as T_1 increases from 0 to 1440.

As for the subtle shape of the curve of f_0^U, it is quite natural that the advantage that the seasonal staffing policy has over the unseasonal one is small when either the peak season or off-peak season is far shorter than the other. As the length ratio of the two seasons moves toward 1 from either side, the advantage increases and finally peaks when $T_0/T_1 = 1$. Thus, the curve of $f_0^* - f_0^U$ should present an inverted U shape (first increases and then decreases). Then the marginally increasing growth trend of the curve of f_0^U with respect to T_1 can be explained by combining the linearity of the curve of f_0^* with respect to T_1 and the inverted U shape of the curve of $f_0^* - f_0^U$ with respect to T_1.

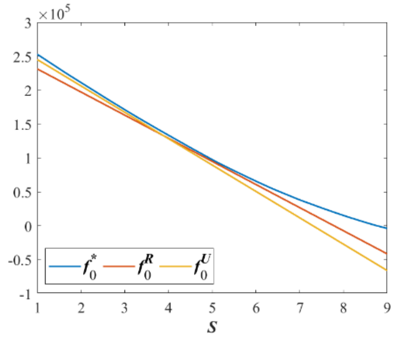


FIGURE 15. Effect of S on f_0^*, f_0^R, and f_0^U.

6) EFFECT OF THE BASIC SALARY OF THE PERMANENT SERVERS (MEASURED BY S)

The trends of f_0^*, f_0^R, and f_0^U with increasing S are shown by Fig.15, in which we can see that: (1) there is always $f_0^* > f_0^R$ and $f_0^* > f_0^U$ for any $S \in [1, 9]$; (2) as S increases, f_0^* decreases with a marginal diminishing effect, while f_0^R and f_0^U decreases linearly; (3) the curve of f_0^R decreases faster than that of f_0^U, and they intersect at a point around $S = 4$.

The downward trends on the curves of f_0^*, f_0^R, and f_0^U with respect to S are expected — rising cost leads to falling profit. The linearity of the curves of f_0^R and f_0^U is also apparent — f_0 is negatively linearly correlated with S while S doesn't affect $\left(\boldsymbol{C}^{\mathbf{R}}, \boldsymbol{\Delta}^{\mathbf{R}}\right)$ and $\boldsymbol{C}^{\mathbf{U}}$ (see equation (23) and (24)). And the faster rate of descent of the curve of f_0^R with respect to S compared to that of f_0^U with respect to S is attributed to a larger number of servers under the rough seasonal staffing method than the unseasonal one.

An interesting trend is seen on the curve of f_0^* with respect to S. The rate of descent of the curve gradually slows down, which leads to an increasing advantage of our method over the other two traditional methods when S increases at a high level. This is because our method is based on a profit maximizing model M_1 so that this method will automatically adjust the personnel allocation to mitigate the increase of costs caused by increasing salaries. The evidence of this explanation can be found in Fig.16, in which there is an obvious downtrend on c_5^* as well as $c_5^* + \delta_5^*$. This effect of reducing cost of our method also reflects its high flexibility of human resource allocation and high performance.

7) EFFECTS OF OTHER PARAMETERS

The effects of the other parameters are relatively simpler and less important, so we do not present them in detail. Instead, we give a brief description.

(1) The curves of f_0^*, f_0^R, and f_0^U with respect to N_5 increase with marginal diminishing effects, because a rise of N_5 reduces the rate of balking and reneging, but it is a waste of space to have a too large customer waiting area. In fact, N_5 concerns another physical resource — customer waiting area in the business hall, which may also have limitation on the effects of λ_5, σ_5, and R_5, and may interact with W_5 and H.

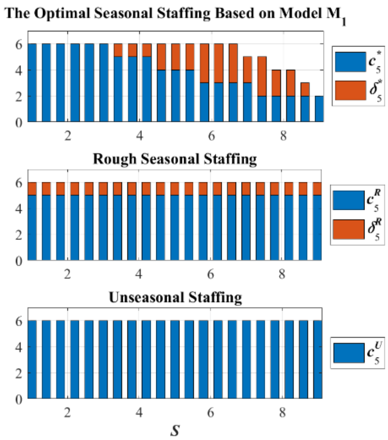


FIGURE 16. Effect of S on $\left(c_i^*, \delta_i^*\right)$, $\left(c_i^R, \delta_i^R\right)$, and c_i^U.

(2) Parameters α presents linearly negative impacts on f_0^*, f_0^R, and f_0^U, because f_0 is negatively linearly correlated with α while α hardly affects the three staffing policies.

(3) Parameters β presents linearly positive impacts on f_0^*, f_0^R, and f_0^U, because f_0 is positively linearly correlated with β while β hardly affects the three staffing policies.

(4) Parameters γ presents linearly negative impacts on f_0^*, f_0^R, and f_0^U, because f_0 is negatively linearly correlated with γ while γ hardly affects the three staffing policies.

(5) Parameters ζ presents linearly negative impacts on f_0^R and f_0^U, and marginally diminishing negative impacts on f_0^*. The explanations for this are similar with that provided in the analyses on the effects of S.

D. DISCUSSION

In the above numerical experiments, we test the effects that the parameters in Θ have on f_0^*, f_0^R, and f_0^U and the three staffing policies. We mainly present and analyze the effects of parameters λ_5, σ_5, W_5, H, T_1, and S. Our results mainly involve two aspects.

Firstly, as long as the season demand variation exists, our seasonal staffing method based on model M_1 always has better performance (measured by the profit) than the rough seasonal staffing method and the unseasonal staffing method (two traditional ways of staffing used by most of banks). This advantage is attributed to our precise and reasonable optimization model that results in the higher efficiency and flexibility in allocating human resource.

However, although the advantage of our method always exists, its efficiency and flexibility may be restricted when the available human resource or the physical resource is scarce. And the restrictions of different resources interact with each other. Such restriction and interaction further indicate the importance of resources integration. Actually, customers, services, and different kinds of resources form a complicated network in which each kind of resource serves as a support for the services that are provided to meet the customer demand; any two kinds of resources may have complex interactions. It is apparently useful for improving the efficiency of resource allocation to analyse the problems of resource interaction and integration [51]–[53].

The second aspect of our results includes a series of results regarding the effects of each parameter on the profit and staffing policy under our seasonal staffing method based on model M_1. The main results are summarized as follow:

1) The customer demand, the seasonal demand variation and the transaction revenue have marginally diminishing positive effects on the profit under our method. The marginal diminishing effects here are attributed to the restriction of human resource or other kinds of resources.
2) The number of service windows, the service system capacity, and the number of staff members available for transferring in peak season all have marginally diminishing positive effects on the profit in certain ranges, and have no impacts when the corresponding parameters are beyond the ranges. And these three parameters interact each other — when any of them is at a low level, the upper bounds of the other two will fall.
3) There is a directly proportional relationship between the profit under our method and the length of the peak season.
4) The profit under our method decreases with a marginal decreasing effect as the basic salary rises. The marginal decreasing effect here indicates a mitigative effect of our method on the reduction of profit brought by an increase of costs. This also indicates the high efficiency of our method in allocating human resource.

V. CONCLUSION

The seasonal demand variation faced by commercial banks has attracted much attention from the industry because of its significant influence on the bank's operation. It is necessary for the bank to implement some seasonal resource allocation policies if the bank wishes to avoid both shortages and surpluses of resources under a significant seasonal demand variation. As a key resource for the commercial bank, human resource deserves special attention. However, because of a lack of in-depth research on the issue of seasonal human resource allocation (staffing), the bank can only design the seasonal staffing policy based on a rough estimation of customer demands in the peak season and off-peak season. Consequently, overstaffing and understaffing occur frequently and cause damage to the bank's operation.

This paper develops a discrete profit maximizing model to determine the optimal numbers of servers for each branch of the bank in peak season and off-peak season. We use a n-dimensional M/M/c/N queueing system with balking and reneging to model the service systems in the n branches. Considering balking and reneging phenomena make our model close to reality. And the n-dimensional model enables our model to be more general and applicable since it is capable of finding a global optimization of human resource allocation in the bank with n branches (n can be any positive integer).

Through a process of linearization, we simplify the model so that a numerical solution can be obtained in a short time period with the help of a computer program. Then, a detailed procedure for applying our model to practice is given as a guidance of applying our model-based seasonal staffing method.

Finally, we conduct a group of numerical experiments. The results verify the advantage of our model-based seasonal staffing method over the traditional methods, including the rough seasonal staffing method and the unseasonal staffing method. Moreover, the numerical results show some interesting effects that some key factors have on the performance and advantage of our seasonal staffing method. And the results also reveal some deeper managerial implications regarding the issue of resource interaction and integration.

This research still has some limitations and can be extended in the future research. Firstly, we do not consider time difference of peak or off-peak season in different branches (for instance, branch B_1 is in peak season while B_2 is in off-peak season). Secondly, we only divide a year into two seasons (peak season and off-peak season), but the situation might be more complex in reality (for instance, a year consists of a peak season, an off-peak season and another intervening season). Thirdly, we just do some numerical analyses based on the data that is not enough to support empirical research. If more data can be collected, an empirical study can be done to enrich the results.

ACKNOWLEDGMENT

The authors would like to thank Prof. Pinedo in the Stern School of Business at New York University for his valuable comments and suggestions. And they also would like to thank the editors of American Journal Experts for helping them with improving English language, grammar, punctuation, spelling, and overall style.

REFERENCES

[1] C. T. Taylor. (1963). *Meeting Seasonal Loan Demands: A Problem of Managing Bank Funds*. [Online]. Available: https://fraser.stlouisfed.org/files/docs/publications/frbatlreview/pages/64293_1960-1964.pdf

[2] INVESTOPEDIA. (2015). *Is the Banking Sector Subject to any Seasonal Trends?* [Online]. Available: https://www.investopedia.com/ask/answers/052915/banking-sector-subject-any-seasonal-trends.asp

[3] Aldermore. *Managing Seasonal Demand*. Accessed: 2019. [Online]. Available: https://www.aldermore.co.uk/business/challenges/managing-seasonal-demand/

[4] E. Petridou and N. Glaveli, "Human resource development in a challenging financial environment: The case of a Greek bank," *Hum. Resource Develop. Int.*, vol. 6, no. 4, pp. 547–558, 2003.

[5] E. D. Hatzakis, S. K. Nair, and M. Pinedo, "Operations in financial services—An overview," *Prod. Oper. Manage.*, vol. 19, no. 6, pp. 633–664, 2010.

[6] A. K. M. Masum, M. A. K. Azad, and L.-S. Beh, "The role of human resource management practices in bank performance," *Total Qual. Manage. Bus. Excellence*, vol. 27, nos. 3–4, pp. 382–397, 2016.

[7] B. Hansen, "Excess demand, unemployment, vacancies, and wages," *Quart. J. Econ.*, vol. 84, no. 1, pp. 1–23, 2006.

[8] A. Parisio and C. N. Jones, "A two-stage stochastic programming approach to employee scheduling in retail outlets with uncertain demand," *Omega*, vol. 53, pp. 97–103, Jun. 2015.

[9] V. Mani, S. Kesavan, and J. M. Swaminathan, "Estimating the impact of understaffing on sales and profitability in retail stores," *Prod. Oper. Manage.*, vol. 24, no. 2, pp. 201–218, 2015.

[10] A. Brinkhoff, Ö. Özer, and G. Sargut, "All you need is trust? An examination of inter-organizational supply chain projects," *Prod. Oper. Manage.*, vol. 24, pp. 181–200, Feb. 2015.

[11] P. Afèche, M. Araghi, and O. Baron, "Customer acquisition, retention, and service access quality: Optimal advertising, capacity level, and capacity allocation," *Manuf. Service Oper. Manage.*, vol. 19, no. 4, pp. 674–691, 2017.

[12] G. D. Quirin and J. L. Bower, "Managing the resource allocation process: A study of corporate planning and investment," *J. Finance*, vol. 26, no. 1, pp. 208–209, Mar. 1971.

[13] G. T. S. Ho, W. H. Ip, C. K. M. Lee, and W. L. Mou, "Customer grouping for better resources allocation using GA based clustering technique," *Expert Syst. Appl.*, vol. 39, pp. 1979–1987, Feb. 2012.

[14] H. Mo, M. Xie, and G. Levitin, "Optimal resource distribution between protection and redundancy considering the time and uncertainties of attacks," *Eur. J. Oper. Res.*, vol. 243, pp. 200–210, May 2015.

[15] F. V. Mitsakis, "Employees' perspectives on strategic human resource development before and after the global financial crisis: Evidence from the Greek banking sector," *Int. J. Training Develop.*, vol. 21, no. 4, pp. 285–303, 2017.

[16] M. Faig, "Seasonal fluctuations and the demand for money," *Quart. J. Econ.*, vol. 104, no. 4, pp. 847–861, 2006.

[17] J. C. Arismendi, J. Back, M. Prokopczuk, R. Paschke, and M. Rudolf, "Seasonal stochastic volatility: Implications for the pricing of commodity options," *J. Banking Finance*, vol. 66, pp. 53–65, May 2016.

[18] Á. Cartea and M. G. Figueroa, "Pricing in electricity markets: A mean reverting jump diffusion model with seasonality," *Appl. Math. Finance*, vol. 12, no. 4, pp. 313–335, 2005.

[19] M. J. Kamstra, L. A. Kramer, and M. D. Levi, "A careful re-examination of seasonality in international stock markets: Comment on sentiment and stock returns," *J. Banking Finance*, vol. 36, no. 4, pp. 934–956, 2012.

[20] T. Y. Chang, S. M. Hartzmark, D. H. Solomon, and E. F. Soltes, "Being surprised by the unsurprising: Earnings seasonality and stock returns," *Rev. Financial Stud.*, vol. 30, no. 1, pp. 281–323, 2017.

[21] S. Anily and M. Haviv, "Cooperation in service systems," *Oper. Res.*, vol. 58, no. 3, pp. 660–673, 2009.

[22] F. Karsten, M. Slikker, and G.-J. van Houtum, "Resource pooling and cost allocation among independent service providers," *Oper. Res.*, vol. 63, no. 2, pp. 476–488, 2015.

[23] Y. Yu, S. Benjaafar, and Y. Gerchak, "Capacity sharing and cost allocation among independent firms with congestion," *Prod. Oper. Manage.*, vol. 24, no. 8, pp. 1285–1310, 2015.

[24] J. M. Calabrese, "Optimal workload allocation in open networks of multiserver queues," *Manage. Sci.*, vol. 38, no. 12, pp. 1792–1802, 2008.

[25] A. J. Rolfe, "A note on marginal allocation in multiple-server service systems," *Manage. Sci.*, vol. 17, no. 9, pp. 656–658, 2008.

[26] M. E. Dyer and L. G. Proll, "Note—On the validity of marginal analysis for allocating servers in *M/M/c* queues," *Manage. Sci.*, vol. 23, no. 9, pp. 1019–1022, 2008.

[27] C. J. Ancker, Jr., and A. V. Gafarian, "Some queuing problems with balking and reneging. I," *Oper. Res.*, vol. 11, no. 1, pp. 88–100, 1963.

[28] C. J. Ancker, Jr., and A. V. Gafarian, "Some queuing problems with balking and reneging—II," *Oper. Res.*, vol. 11, no. 6, pp. 928–937, 1963.

[29] M. O. Abou-El-Ata and A. M. A. Hariri, "The *M/M/c/N* queue with balking and reneging," *Comput. Oper. Res.*, vol. 19, pp. 713–716, Nov. 1992.

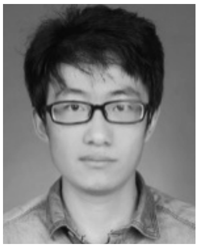
[30] R. O. Al-Seedy, A. A. El-Sherbiny, S. A. El-Shehawy, and S. I. Ammar, "Transient solution of the *M/M/c* queue with balking and reneging," *Comput. Math. Appl.*, vol. 57, pp. 1280–1285, Apr. 2009.

[31] D. Guha, V. Goswami, and A. D. Banik, "Algorithmic computation of steady-state probabilities in an almost observable *GI/M/c* queue with or without vacations under state dependent balking and reneging," *Appl. Math. Model.*, vol. 40, pp. 4199–4219, Mar. 2016.

[32] M. F. Yassen and A. M. K. Tarabia, "Transient analysis of Markovian queueing system with balking and reneging subject to catastrophes and server failures," *Appl. Math. Inf. Sci.*, vol. 11, no. 4, pp. 1041–1047, 2017.

[33] L. Brown, N. Gans, A. Mandelbaum, A. Sakov, H. Shen, S. Zeltyn, and L. Zhao, "Statistical analysis of a telephone call center: A queueing-science perspective," *J. Amer. Stat. Assoc.*, vol. 100, no. 469, pp. 35–50, 2005.

[34] Z. Aksin, M. Armony, and V. Mehrotra, "The modern call center: A multi-disciplinary perspective on operations management research," *Prod. Oper. Manage.*, vol. 16, no. 6, pp. 665–688, 2007.
[35] S. M. R. Iravani, B. Kolfal, and M. P. Van Oyen, "Call-center labor cross-training: It's a small world after all," *Manage. Sci.*, vol. 53, no. 7, pp. 1102–1112, 2007.
[36] D. Marsden and R. Richardson, "Performing for pay? The effects of 'merit pay' on motivation in a public service," *Brit. J. Ind. Relations*, vol. 32, no. 2, pp. 243–261, 1994.
[37] S. J. Procter, L. McArdle, M. Rowlinson, P. Forrester, and J. Hassard, "Performance related pay in operation: A case study from the electronics industry," *Hum. Resource Manage. J.*, vol. 3, no. 4, pp. 60–74, 1993.
[38] E. P. Lazear, "Performance pay and productivity," *Amer. Econ. Rev.*, vol. 90, no. 5, pp. 1346–1361, 2000.
[39] T. Dohmen and A. Falk, "Performance pay and multidimensional sorting: Productivity, preferences, and gender," *Amer. Econ. Rev.*, vol. 101, no. 2, pp. 556–590, 2011.
[40] F. Ederer and G. Manso, "Is pay for performance detrimental to innovation?" *Manage. Sci.*, vol. 59, no. 7, pp. 1496–1513, 2013.
[41] I. M. Toke, "The order book as a queueing system: Average depth and influence of the size of limit orders," *Quant. Finance*, vol. 15, no. 5, pp. 795–808, 2015.
[42] J. N. Tsitsiklis and K. Xu, "Flexible queueing architectures," *Oper. Res.*, vol. 65, no. 5, pp. 1398–1413, 2017.
[43] H. Nazerzadeh and R. S. Randhawa, "Near-optimality of coarse service grades for customer differentiation in queueing systems," *Prod. Oper. Manage.*, vol. 27, no. 3, pp. 578–595, 2018.
[44] H. D. Sherali and W. P. Adams, "A reformulation-linearization technique for solving discrete and continuous nonconvex problems," in *Computers & Mathematics With Applications*. Boston, MA, USA: Springer, 2003.
[45] H. D. Sherali and W. P. Adams, "Reformulation-linearization techniques for discrete optimization problems," in *Handbook of Combinatorial Optimization*. Boston, MA, USA: Springer, 2013.
[46] E. L. Lawler and D. E. Wood, "Branch-and-bound methods: A survey," *Oper. Res.*, vol. 14, no. 4, pp. 699–719, 2008.
[47] I. Griva, S. G. Nash, and A. Sofer, *Linear and Nonlinear Optimization*. Philadelphia, PA, USA: SIAM, 2011.
[48] I. Charon and O. Hudry, "Branch-and-bound methods," in *Concepts of Combinatorial Optimization*, 2nd ed. Hoboken, NJ, USA: Wiley, 2014.
[49] E. Adida and V. DeMiguel, "Supply chain competition with multiple manufacturers and retailers," *Oper. Res.*, vol. 59, no. 1, pp. 156–172, 2011.
[50] Y. Alan and V. Gaur, "Operational investment and capital structure under asset-based lending," *Manuf. Service Oper. Manage.*, vol. 20, no. 4, pp. 637–654, 2018.
[51] E. Gummesson and C. Mele, "Marketing as value co-creation through network interaction and resource integration," *J. Bus. Market Manage.*, vol. 4, no. 4, pp. 181–198, 2010.
[52] E. Baraldi, E. Gressetvold, and D. Harrison, "Resource interaction in inter-organizational networks: Introduction to the special issue," *J. Bus. Res.*, vol. 65, no. 2, pp. 123–127, 2012.
[53] K. Koskela-Huotari, B. Edvardsson, J. M. Jonas, D. Sörhammar, and L. Witell, "Innovation in service ecosystems—Breaking, making, and maintaining institutionalized rules of resource integration," *J. Bus. Res.*, vol. 69, no. 8, pp. 2964–2971, 2016.
[54] D. L. Massart, B. G. M. Vandeginste, S. N. Deming, Y. Michotte, and L. Kaufman, "Operations research," in *Data Handling in Science and Technology*. Amsterdam, The Netherlands: Elsevier, 2003.
[55] L. E. Clarke, D. Gross, and C. M. Harris, "Fundamentals of queueing theory," in *The Mathematical Gazette*. Leicester, U.K.: The Mathematical Association, 2007.
[56] H. G. Heneman and D. W. Belcher, "Wage and salary administration," in *Industrial and Labor Relations Review*. Thousand Oaks, CA, USA: Sage, 2006.



YUANQIAO CHEN was born in China, in 1991. He is currently pursuing the Ph.D. degree with the School of Management and Engineering, Nanjing University, China. He is also working as a Research Fellow with the Urban Finance Research Institute, Industrial and Commercial Bank of China. From September 2017 to September 2018, he spent the academic year with the Stern School of Business, New University, USA, as a Visiting Ph.D. Student. His research interests include the management science and engineering, finance, banking, numerical analysis, and statistics.

XIAOQIANG ZHANG was born in China, in 1991. He received the Ph.D. degree from the School of Management and Engineering, Nanjing University, China, in June 2019, where he is doing a research project with the School of Management and Engineering. He is going to work as a Research Fellow with Shanghai Stock Exchange, China, in September 2019. His research interests include the management science and engineering, finance, and machine learning.

BOMIN BIAN was born in China, in 1990. He is currently pursuing the Ph.D. degree with the Institute of Systems Engineering, College of Management and Economics, Tianjin University, China. From September 2017 and October 2018, he spent the academic year with the Stern School of Business, New University, USA, as a Visiting Ph.D. Student. His research interests include the management science and engineering, finance, and machine learning. His research interests include the management science and engineering, operations research, and transportation.

HAOHUA LI was born in China, in 1990. He received the Ph.D. degree from the School of Management and Engineering, Nanjing University, China, in June 2019, where he is currently working as a Research Fellow with the School of Management and Engineering. His research interests include the management science and engineering, finance, and banking.