Math. Model. Nat. Phenom. 19 (2024) 4
https://doi.org/10.1051/mmnp/2024002

Mathematical Modelling of Natural Phenomena
www.mmnp-journal.org

OPTIMAL ORDERING STRATEGY AND BUDGET ALLOCATION FOR THE COVID-19 VACCINATION PLANNING

Xueping Liu[1], Sheng Zhu[1,*] and Jinting Wang[2]

Abstract. During the COVID-19 pandemic, the most important thing was to control the overall infection rate. To achieve this goal, social managers can choose to use vaccines with different production cycles and therapeutic effects for epidemic prevention and control under financial budget constraints. In this paper we adopt a two-tier queueing system with reneging to characterize the operation management of COVID-19 vaccine ordering and vaccination, in which a higher-efficacy vaccine queue (HQ) and a lower-efficacy vaccine queue (LQ) are employed to account for two types of vaccines service. In light of this framework, a recursive formula is proposed for deriving the infection rates of residents in both HQ and LQ. Social managers can achieve the lowest total infection rate by selecting appropriate vaccine ordering strategies under fixed service capacity, or by allocating financial budgets reasonably under the investment cost regime. Accordingly, we obtain the socially optimal vaccine ordering strategies and financial budget allocation. Finally, we analyze the sensitivity of various parameters to relevant optimal strategies and discover that utilizing a mixed ordering strategy is socially optimal in most circumstances. However, in some extreme cases, ordering a single type of vaccine (higher- or lower-efficacy) may also result in the lowest societal infection rate.

Mathematics Subject Classification. 60K25, 90B22.

Received June 9, 2023. Accepted February 7, 2024.

1. Introduction

The COVID-19 pandemic has resulted in significant loss of life globally, with hundreds of thousands of people succumbing to the disease every month. Vaccination is currently the most effective way to combat the spread of the disease. To minimize losses, especially in the early and middle phases of the pandemic, it is crucial to minimize overall infection rates worldwide. For social managers, including public organizations and governments, one of the things that is challenging is making decisions on the optimal vaccine ordering strategy when faced with shortages of vaccine supply and financial budget constraints.

The progress of biomedical technology is closely following the emergence and development of the epidemic. Based on the study of Castillo *et al.* [1], the vaccines can be classified into two categories on the market, higher-efficacy vaccine and lower-efficacy vaccine. Vaccines with lower efficacy are known for their faster development, lower storage requirements, and shorter manufacturing cycles as compared to vaccines with higher efficacy. For

Keywords and phrases: Queueing system, COVID-19 vaccine, ordering strategy, financial resources allocation.

[1] School of Mathematics and Information Science, Henan Polytechnic University, Henan, PR China.
[2] Department of Management Science, School of Management Science and Engineering, Central University of Finance and Economics Beijing, PR China.

[*] Corresponding author: shengzhu_ms@sina.com

© The authors. Published by EDP Sciences, 2024

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

instance, studies [2] have shown that the mRNA vaccines developed by Pfizer-BioNTech and Moderna have been highly effective in preventing SARS-CoV-2 infection, with efficacy rates of 96% and 98%, respectively. The Pfizer-BioNTech and Moderna vaccines require storage at –80°C and –20°C, respectively, and have a shelf life of up to 6 months (see the work of Uddin and Roni [3]). This highlights the significant financial and logistical resources required for their proper storage. CureVac developed another mRNA vaccine called CVnCoV, which can be stored at higher temperatures (5°C) for at least three months. However, the vaccine candidate used a smaller dose (12 μg) than Pfizer's (30 μg) and showed only 47% efficacy in the phase 2b/3 trial [4]. There is a trade-off between good customer service and high service quality. Good customer service is associated with shorter wait times and queue lengths, while high service quality is linked with longer wait times and queues. Based on the work of Castillo *et al.* [1], it can be concluded that during a pandemic, it may be more beneficial to use a currently available vaccine with lower efficacy rather than waiting for a more effective one due to the importance of speed in vaccine distribution. In light of the fact that COVID-19 viruses mutate over time and that infection rates differ among viral types, it is important to continually reassess and adjust our vaccine ordering strategy. To achieve this, a model should be developed to determine the optimal ordering strategy for vaccines with different efficacy rates.

In this paper, our goal is to minimize the overall infection rate in society by determining the optimal ordering strategies for higher-efficacy and lower-efficacy vaccines or by implementing a reasonable allocation of financial resources. We propose dividing vaccinated residents into two groups: the first group receives higher-efficacy vaccine, while the second group receives lower-efficacy vaccine. Due to the high infectivity of the virus, residents may contract it while waiting in line. In the event that a resident is found to be infected, it is imperative that they are immediately removed from the queue and quarantined. Based on the above characteristics, we can view this model as a two-tier queueing system with reneging. Our primary objective is to maximize social welfare by minimizing the overall infection rate.

It should be noted that although we may currently be in a post-pandemic period, it is important to note that infectious disease crises occur frequently. Throughout history, major outbreaks of diseases such as the plague, cholera, and tuberculosis, among others, have been recorded. In fact, smallpox was the first infectious disease to be listed by the World Health Organization (WHO) as needing to be controlled, back in 1948. Plague has experienced three major outbreaks in history, while cholera has experienced seven outbreaks since 1817. According to the statistics of the World Health Organization, 8 million to 10 million people worldwide suffer from tuberculosis every year. Tuberculosis is currently the deadliest disease, and the death rate of SARS in 2003 was 11%. Our proposed model has broad applicability including but not limited to COVID-19, and can be used to address other infectious diseases that affect humans. As a result, our study has significant implications. Our contributions are outlined below.

- Methodologically, a new formulation of a two-tier queueing system with reneging is proposed to characterize the operation of COVID-19 vaccine ordering and vaccination. A recursive algorithm is designed to explore the infection rates of residents in both the higher-efficacy vaccine queue (HQ) and the lower-efficacy vaccine queue (LQ) during the waiting process.
- Under fixed service capacity, it is revealed that social managers can achieve the lowest infection rate by selecting an appropriate vaccine ordering strategy, with which social losses are greatly reduced. Under the investment cost regime, social managers can achieve the lowest overall infection rate in society through optimal allocation of the financial budget. The socially optimal allocation strategy is obtained.
- Numerically, we investigate the sensitivity of different parameters to relevant optimal strategies and find that, in most cases, using a mixed ordering strategy is socially optimal. However, in some extreme situations, ordering a single type of vaccine (higher- or lower-efficacy) may also achieve the lowest social infection rate.

2. Literature review

Since March 2020, the world has been grappling with a global health crisis caused by the outbreak of a novel coronavirus, resulting in a respiratory disease known as COVID-19 [5]. As our understanding of the virus has evolved, there is mounting evidence that it is more infectious than initially believed [6]. Consequently, scholars have carried out many studies to explore issues related to COVID-19, especially focusing on developing effective prevention and control strategies. Rozhkov *et al.* [7] investigated the impact of the COVID-19 outbreak on network design structures and optimized supply chain management by making adaptive operational decisions. Similarly, intending to reduce the risk of viral transmission during patient transfers, Hosseini-Motlagh *et al.* [8] presented a control strategy to rationally assign COVID-19 patients to screening facilities, healthcare facilities, and quarantine facilities in order to minimize the transmission of the virus. The results of the study suggest that assigning patients to appropriate healthcare centers can effectively improve the efficiency of the healthcare system. In addition, many scholars have conducted in-depth research from multiple perspectives using queueing theory, aiming to reduce the risk of residents being infected with COVID-19. In order to minimize the risk of infection among residents waiting in line at department stores and supermarkets, business managers must limit the number of shoppers present at any given time. Perlman and Yechiali [9] developed two models to enhance security by optimizing queue size and minimizing customer's waiting time. Then, to ensure minimal infection rates, they used the Stackelberg game to determine an equilibrium strategy for the maximum number of shoppers allowed in the store. The findings of this study have significant implications for both authorities and companies in terms of protecting the safety of customers and employees, as well as reducing costs. Concerning the scheduling optimization of COVID-19 testing resources, Yang *et al.* [10] developed a queueing game-theoretic model, which simulates the queueing process of residents taking COVID-19 tests. By optimizing the schedule and price of testing resources, the model successfully helped to identify more infected cases, which made an important contribution to the epidemic prevention and control efforts. Then, Kang *et al.* [11] introduced a new modeling framework for assessing the risk of disease transmission in small-scale environments during pandemics. They developed a queueing system with overlapping sojourn times and derived performance measures for queueing models under different service rules (first-come-first-served or priority service for specific groups). Furthermore, the public demand for vaccination is extremely urgent during virus pandemics. However, if the vaccine brand is not the preferred choice of the residents, they may refuse to be vaccinated. To utilize resources efficiently, policymakers may choose not to disclose information about vaccine brands, thereby limiting residents' choices. Nageswaran [12] developed a queueing game model to assess the effectiveness of this measure and analyzed the dynamic equilibrium relationship between individuals and policymakers. The results of the study show that restricting individuals' choices can motivate more people to vaccinate in the case of vaccine shortages. Therefore, in this paper, we also adopt the queueing game model to solve the COVID-19 vaccination and ordering problem.

According to a study on vaccine supply (see the work of Castillo *et al.* [1]), widespread immunization just one month earlier could result in saving numerous lives and mitigating both short- and long-term economic damage to the government. Thus, achieving community immunity by vaccinating the entire population as soon as possible is the only way to curb the spread of the virus. Tang *et al.* [13] employed a bi-objective mixed integer linear programming (MILP) model to optimize both the level of service and operating costs of vaccinators. The researchers achieved this by controlling the total travel distance of vaccinators and the total cost. The results show that their approach reduces operating costs and total travel distance by 9.3% and 36.6%, respectively. The management implications suggest that expanding the service capacity of vaccination sites can help improve the performance of vaccination programs. What's more, unlike optimizing the number and location of vaccination sites, achieving community immunity can also be accelerated by standardizing the pricing of the COVID-19 vaccine. Martonosi *et al.* [14] modeled the U.S. COVID-19 vaccine market as a duopoly between two manufacturers, Pfizer-Biotech and Moderna, and used the optimal game-theoretic approach, which was proposed by Cummings *et al.* [15], to determine the price of COVID-19 vaccines. Their work further increased the willingness of the public to receive the vaccine. However, while there has been significant research on vaccine pricing and the optimal location of vaccination sites, little attention has been given to the topic of vaccine ordering.

Ordering too few vaccines may not control the spread of the COVID-19 virus, while ordering too much may lead to a waste of resources. In this paper, we divide the vaccines in the market into two types, *i.e.*, higher-efficacy and lower-efficacy vaccines, and use a two-tier queueing model to investigate the resource allocation problem for vaccines, so as to determine the optimal ordering ratio for the two types of vaccines. Our research involves queueing models with reneging customers, two-tier mechanisms, and service resource allocation. Next, we will conduct a literature review on these three aspects.

• **Queueing models with reneging customers**. In the context of COVID-19, it is crucial to transfer infected residents promptly during the waiting process. To address this issue, we will develop a queueing model with reneging customers. Barrer [16] first studied the queueing problem with impatient customers, where customers are randomly served while waiting. He obtained an equilibrium distribution of queue length and calculated the average exit rate in the queueing system. Subsequently, Barrer [17] expanded his research to address the issue of impatient customers waiting in a "first-come, first-served" queue. He focused on the ratio of the average customer "lost" rate to the average arrival rate as the main parameter of interest. Through statistical equilibrium, he derived closed expressions for these ratios and determined the distribution of queue lengths. Ancker and Gafarian [18] studied the M/M/1 queueing model with balking and reneging customers. They obtained performance measures such as the average queue length and the reneging rate for this model in the steady state situation. Besides, Sarhangian and Balcioglu [19] studied the priority queueing system with impatient customers, in which the patience of customers obeys a negative exponential distribution. They derived the Laplace transform of the virtual waiting time for each case using the level-crossing method and calculated some performance measures. Interested readers may also refer to the work of Economou *et al.* [20] for specific details on queueing systems with reneging customers.

• **Two-tier queueing models**. To model the interaction between the two vaccination channels, we create a two-tier queueing network that includes two types of vaccination services. One service has high quality but low speed, while the other has low quality but high speed. This two-tier queueing model allows residents to choose the queue that best fits their needs when they decide to join the system. Two-tier service systems were first proposed by Tuohy *et al.* [21]. Subsequently, it was extensively studied by scholars due to its practical applications. One such application is the border crossings between the United States and Canada, where the issue of waiting in line is a typical two-tier queueing pattern. To aid inspectors in identifying suspicious individuals, it is necessary to maintain a queue length that allows for a 10–15 min waiting time for arriving vehicles. To ensure a positive customer service experience, the system aims to maintain a reasonable queue length by controlling the number of open traffic stops. To achieve this, Zhang [22] introduced a Markov-based benchmark model that utilizes geometric solutions to optimize the number of open traffic stations. Furthermore, in a hybrid cloud system, there exists a two-tier model where web service companies requiring computing resources allocate some of their upcoming work to the cloud and the remainder to their on-premise subsystems to avail services. Considering the strategic behavior of customers, Zhu *et al.* [23] studied the optimal control in this hybrid cloud system with a two-tier mechanism. They analyzed customers' behavior strategies and obtained the optimal admission control approach to maximize the server profit and social welfare. Other works on this topic include Guo *et al.* [24], Hua *et al.* [25], and others.

• **Resource allocation problem**. In this paper, we explore the resource allocation problem of minimizing the overall infection rate by determining the optimal proportion of higher and lower efficacy vaccines. This issue is also present in other healthcare systems. In practice, both general hospitals (*e.g.*, AAA hospitals) and primary hospitals (*e.g.*, community hospitals) exist in the healthcare system. Wang *et al.* [26] developed a queueing game theoretical model to solve the resource allocation problem in a public service system with multiple types of providers and patients. The model aims to address how limited resources can be allocated to both types of hospitals to maximize social welfare. Through this approach, the authors obtained a socially optimal reimbursement policy. Although the optimal solution for agents may result in a heavily imbalanced resource allocation, the equilibrium allocation may also not be the most efficient. To address this issue, Nicosia *et al.* [27] evaluated the quality of the equitable solution by creating a model that measured the loss of system efficiency under equitable allocation. Furthermore, effective allocation of vaccine and treatment resources is crucial for controlling the spread of an epidemic within a fixed budget. Zhang *et al.* [28] developed an ensemble population network model to simulate the spatial progression of an epidemic. They employed a binary particle swarm optimization (BPSO) algorithm to identify the optimal solution for resource deployment. The study found that in situations where the budget is limited, there is a preference for allocating resources towards vaccines rather than treatments. However, as the budget increases, there is a shift towards allocating more resources to treatment options, while also ensuring that vaccine resources remain adequate. What's more, Chen *et al.* [29] proposed a multi-objective optimization model for cloud resource allocation that takes into account the minimum number of physical servers used and the minimum matching distance between resource performance and resource ratio. Their algorithm addresses the issue of traditional cloud resource allocation methods not supporting contingency models, thus ensuring both timeliness and optimization of resource allocation.

The rest of this paper is organized as follows. In Section 3, we provide a detailed model description. In Section 4, we derive recursive formulas for the infection rates of residents in both higher- and lower-efficacy vaccine queues during the waiting process, and then obtain socially optimal ordering strategies. Section 5 gives a socially optimal financial budget allocation under the investment cost regime. Conclusions and future works are provided in Section 6.

3. Model description

Consider a vaccine ordering system consisting of two types of providers, in which residents who need to be vaccinated arrive according to a Poisson process with rate λ. The service discipline is on a "first come, first served" basis. This system has two types of vaccines: higher-efficacy vaccine and lower-efficacy vaccine. There are two queues in the system, namely the higher-efficacy vaccine queue (HQ) and the lower-efficacy vaccine queue (LQ). A fixed percentage of residents will join the HQ, while the remainder will join the LQ. Residents who join the HQ can obtain vaccines with a higher potency, whereas those who join the LQ can only receive vaccines with a lower potency. With the increase in waiting time, waiting residents may be infected. We assume that residents are infected with the rate i_h in the HQ and with rate i_l in the LQ. Because infected residents may have adverse effects on other waiting residents and do not require vaccinations, they will be withdrawn from the queues. Therefore, the assumption that reneging is allowed is reasonable, and it is clearly in line with the actual situation.

However, it is not always possible for residents to be vaccinated immediately after joining the HQ or LQ. Not only do residents have to wait for the vaccine to be produced, transported, and finally delivered to the community service station, but they also need to wait for the previous residents to finish the service if the vaccine is available. We assume that T_1 and T_2 denote the service times in the HQ and LQ, respectively. Thus, the service time T_1 (or T_2) is composed of three parts: the time for producing vaccines T_{1a} (or T_{2a}), the time for transporting vaccines T_{1b} (or T_{2b}) and community service time T_{1c} (or T_{2c}), and we assume that they all are exponentially distributed with rate α_1 (or α_2), β_1 (or β_2) and γ, respectively. Then, we can denote the service times T_1 and T_2 as

$$T_i = c_1 T_{ia} + c_2 T_{ib} + c_3 T_{ic}, \quad i = 1, 2, \tag{3.1}$$

where c_1, c_2 and c_3 are constants. If the service times T_i $(i = 1, 2)$ are assumed to be generally distributed with rate μ_i (i=1,2) respectively, then μ_i must satisfy

$$\frac{1}{\mu_i} = c_1 \frac{1}{\alpha_i} + c_2 \frac{1}{\beta_i} + c_3 \frac{1}{\gamma}, \quad i = 1, 2. \tag{3.2}$$

However, residents can still be infected with the disease after they have been vaccinated with a higher- or lower-efficacy vaccine. We use i_{hf} and i_{lf} to represent the infection rate of residents after receiving higher- and lower-efficacy vaccines, respectively.

The above model can be characterized as a two-tier queueing system with reneging residents. Among all arriving residents, a proportion of θ residents join HQ, while the rest join LQ. Then, the effective arrival rate of joining the HQ (or LQ) is $\lambda\theta$ (or $\lambda(1-\theta)$), denoted by λ_1 (or λ_2), namely, $\lambda_1 = \lambda\theta$ and $\lambda_2 = \lambda(1-\theta)$. Furthermore, residents may renege from the queues due to infection. The reneging of residents follows a Poisson process with rate i_h (or i_l) in the HQ (or LQ). The service times are generally distributed, with a high rate μ_1 in the HQ and a low rate μ_2 in LQ, respectively. In addition, the infection rate among residents who have

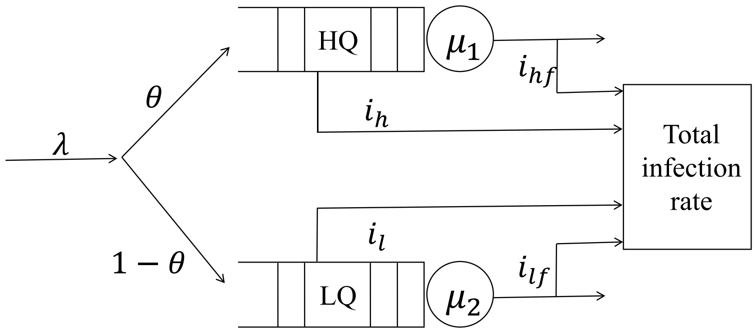


FIGURE 1. Schematic diagram of the vaccine ordering model.

TABLE 1. Notations and definitions.

Notation	Definition
λ	Total arrival rate of residents
λ_1	Arrival rate of residents who join the HQ
λ_2	Arrival rate of residents who join the LQ
μ_1	Service rate in the HQ
μ_2	Service rate in the LQ
θ	The proportion of residents joining the HQ
I_a	Total infection rate of residents
i_h	Infection rate of residents in the HQ
i_l	Infection rate of residents in the LQ
i_{hf}	Infection rate of residents after the higher-efficacy vaccination
i_{lf}	Infection rate of residents after the lower-efficacy vaccination
$f_1(t)$	Probability density function of the service time in the HQ
$f_2(t)$	Probability density function of the service time in the LQ
$F_1(x)$	Probability distribution function of the service time in the HQ
$F_2(x)$	Probability distribution function of the service time in the LQ

completed higher-efficacy (or lower-efficacy) vaccination is i_{hf} (or i_{lf}). In reality, vaccines with higher efficacy are more dependable but have longer wait times, while vaccines with lower efficacy have shorter wait times but are less likely to be effective. A schematic diagram of the vaccine ordering model is shown in Figure 1.

Our overall goal is to find the optimal ordering strategy or the optimal service capacity that minimizes the total infection rate. However, in the context of the COVID-19 pandemic, vaccine resources are extremely scarce and vaccine allocation is determined by the government. It is difficult for residents to be vaccinated according to their own preferences, and the proportion of residents joining different queues depends on the vaccine allocated by the government. Therefore, the proportion of residents joining different queues is also the government's strategy for ordering vaccines in this paper, *i.e.*, θ denotes the proportion of higher-efficacy vaccines ordered by the government and 1-θ denotes the proportion of lower-efficacy vaccines ordered. Accordingly, the service capacity of the system depends on the service time T_i, the longer the service time, the worse the service capacity of the system. From the earlier description, it is clear that the service time is determined by three variables, namely T_{ia}, T_{ib} and T_{ic}, which obey exponential distributions with α_i, β_i and γ, respectively. Thus, α_i, β_i, and γ denote the capacity to produce vaccines, the capacity to transport vaccines, and the community service capacity. The capacity of the system service then depends on the size of α_i, β_i, and γ. Table 1 provides the notations used in this paper.

We have described the queueing model of vaccine ordering and the meaning of each parameter in the modeling process above, and we will explain the rationality of the assumptions related to the parameters in the following:

• **Mechanisms for vaccination services.** In this paper, we categorize the efficacy of vaccines into two types: higher-efficacy vaccines and lower-efficacy vaccines. Compared to lower-efficacy vaccines, higher-efficacy vaccines have stronger protective effects, but require longer waiting times in queues. Residents need to balance effectiveness and waiting time. Therefore, we model the vaccine ordering system as a two-tier queueing model. Our objective is to analyze which vaccine ordering strategies can maximize social welfare from the government's perspective. For social planners, they mainly focus on the number of people vaccinated every day and the overall infection rate, rather than the specific order of individual vaccination. Thus, this paper does not consider the case where priority vaccination is offered to some residents but assumes that residents are vaccinated according to the first-come, first-served (FCFS) mechanism.

• **COVID-19 vaccination and distribution strategies.** Similar studies on COVID-19 vaccination and distribution strategies can be found in the work of Franco *et al.* [30] and Jahani *et al.* [31]. Franco *et al.* [30] used the queueing network to model the medical treatment process of COVID-19, and investigated the arrangement of medical staff and material resources in the vaccination process during the COVID-19 virus pandemic based on the experimental data of Colombia. Jahani *et al.* [31] studied vaccine supply chain management strategies triggered by crises based on real data from Australia. They set up a queueing model to simulate the vaccination process, and distributed different types of COVID-19 vaccines to people with different levels of vulnerability. When modeling the queueing process of vaccination, the above literature assumes that residents arrive at vaccination stations with a Poisson flow. In our work, we also adopt similar assumptions.

• **Vaccination service time.** In our manuscript, we assume that the community vaccination time of residents follows an exponential distribution. This hypothesis follows the work of Franco *et al.* ([30]), who collected data on vaccination service time and found that exponential distribution is very suitable for vaccination and registration time models. In addition, Jahani *et al.* [31] found that the speed of vaccination is influenced by supply chain efficiency, such as vaccine production and transportation time. Therefore, in our manuscript, we assume that the service time is also affected by vaccine production time and transportation time.

• **Statistical analysis of infection rate.** According to the global COVID-19 daily count cases data from the Worldometer's official website[1], we find the data from February 2020-March 2020 to statistically analyze and obtaine Figure 2, we found that the infection rate during a COVID-19 virus pandemic shows an exponential trend over time (see Fig. 2), which embodies the reasonability of our assumption. It is crucial to control the virus as soon as possible during a pandemic. Our study focuses on the optimal vaccine ordering strategy during a pandemic when vaccine resources are in short supply, So it is reasonable to assume that the infection rate shows exponential growth during the pandemic.

4. Service capacity regime

Under the service capability regime, the service capacities of the higher- and lower-efficacy vaccine queueing systems (*i.e.*, μ_1 and μ_2) are fixed, and social managers achieve the lowest infection rate by selecting an appropriate vaccine ordering strategies. In this section, we consider the optimal ordering strategies for the two types of vaccines that minimize the total infection rate. Theorem 4.1 gives the probability density functions of the service times T_1 and T_2. Theorem 4.3 shows the mean waiting time of residents in the HQ and LQ.

4.1. System Performance

In order to reduce society's overall infection rate and swiftly stop the COVID-19 virus's propagation, we must determine the optimal ordering strategies for both vaccine classes. To achieve this goal, we must first explore the relevant performance measures of this system and obtain the overall infection rate in society based on these results.

[1] Data from https://www.worldometers.info/coronavirus/.

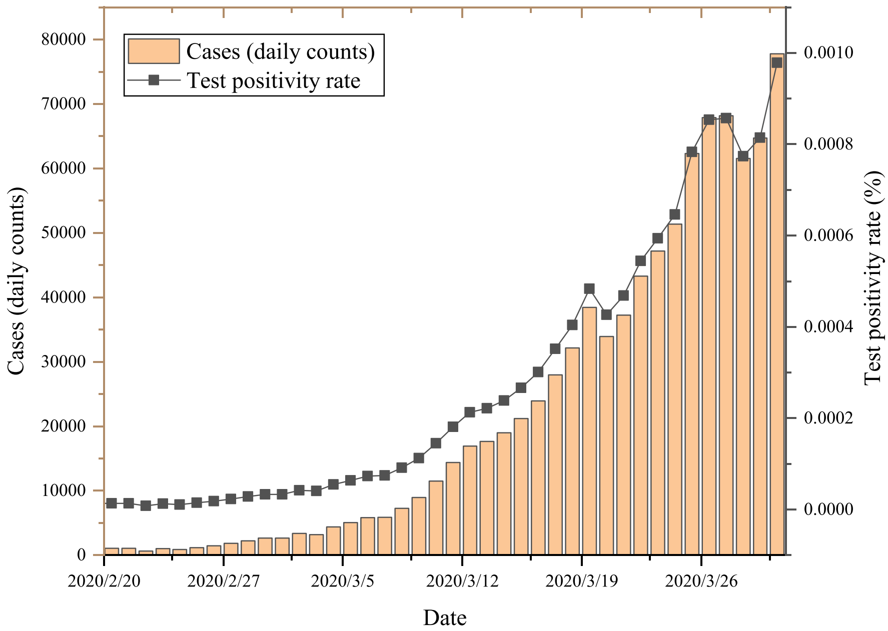


FIGURE 2. Global daily counts of COVID-19 cases and test-positive rates.

Theorem 4.1. *The probability density function of the service times T_i (i=1,2) can be written as*

$$f_i(t)=\begin{cases}\dfrac{\alpha_i\beta_i\gamma c_1(\gamma c_2-\beta_i c_3)\mathrm{e}^{-\frac{\alpha_i}{c_1}t}-\alpha_i\beta_i\gamma c_2(\gamma c_1-\alpha_i c_3)\mathrm{e}^{-\frac{\beta_i}{c_2}t}}{(\beta_i c_1-\alpha_i c_2)(\gamma c_1-\alpha_i c_3)(\gamma c_2-\beta_i c_3)}+\dfrac{\alpha_i\beta_i\gamma c_3\mathrm{e}^{-\frac{\gamma}{c_3}t}}{(\gamma c_1-\alpha_i c_3)(\gamma c_2-\beta_i c_3)} & t>0,\\ 0 & else.\end{cases} \tag{4.1}$$

Proof. Based on the model assumption in Section 3, $T_{ia}\sim Exp(\alpha_i)$, $T_{ib}\sim Exp(\beta_i)$ and $T_{ic}\sim Exp(\gamma)$, $i=1,2$. If $x\le 0$, $P\{c_1T_{1a}\le x\}=P\{T_{1a}\le\frac{x}{c_1}\}=0$. If $x>0$, we obtain

$$P\{c_1T_{1a}\le x\}=P\{T_{1a}\le\frac{x}{c_1}\}=\int_{-\infty}^{\frac{x}{c_1}}f(t)\mathrm{d}t=\int_0^{\frac{x}{c_1}}\alpha_1\mathrm{e}^{-\alpha_1 t}\mathrm{d}t=1-\mathrm{e}^{-\alpha_1\frac{x}{c_1}}.$$

Thus, we can get the probability density function of c_1T_{1a}, denoted by $f_{c_1T_{1a}}(y)$, that is,

$$f_{c_1T_{1a}}(y)=\begin{cases}\dfrac{\alpha_1}{c_1}\mathrm{e}^{-\frac{\alpha_1}{c_1}y} & y>0, c_1>0,\\ 0 & else.\end{cases}$$

Hence, c_1T_{1a} is exponentially distributed with rate $\frac{\alpha_1}{c_1}$. Similarly, we can find that c_1T_{ia}, c_2T_{ib} and c_3T_{ic} $(i=1,2)$ are exponentially distributed with the rates of $\frac{\alpha_i}{c_1}$, $\frac{\beta_i}{c_2}$ and $\frac{\gamma}{c_3}$ (i=1,2), respectively. According to (3.1) and the work of Song [32], we can obtain the probability density functions of T_i $(i=1,2)$ as (4.1). This completes the proof. □

Theorem 4.1 gives the probability density functions of service times T_1 and T_2 in the HQ and LQ. Since the probability distribution function can be obtained by integrating the probability density on $(-\infty, x]$, we can easily obtain the corresponding probability distribution functions of T_1 and T_2. Let $\bar{F}_1(x) = 1 - F_1(x)$ and $\bar{F}_2(x) = 1 - F_2(x)$, where $\bar{F}_1(x)$ denotes the complementary distribution of service time in the HQ and $\bar{F}_2(x)$ denotes the complementary distribution of service time in the LQ. We summarize the related results in the following corollary.

Corollary 4.2. *The probability distribution functions and the complementary probability distributions of* $T_i, i = 1, 2,$ *satisfy:*

$$F_i(x) = 1 - u_i \mathrm{e}^{-\frac{\alpha_i}{c_1}x} + v_i \mathrm{e}^{-\frac{\beta_i}{c_2}x} - w_i \mathrm{e}^{-\frac{\gamma}{c_3}x}, \quad i = 1, 2, \tag{4.2}$$

and

$$\bar{F}_i(x) = u_i \mathrm{e}^{-\frac{\alpha_i}{c_1}x} - v_i \mathrm{e}^{-\frac{\beta_i}{c_2}x} + w_i \mathrm{e}^{-\frac{\gamma}{c_3}x}, \quad i = 1, 2, \tag{4.3}$$

where $u_i = \frac{\beta_i \gamma c_1^2}{(\beta_i c_1 - \alpha_i c_2)(\gamma c_1 - \alpha_i c_3)}$, $v_i = \frac{\alpha_i \gamma c_2^2}{(\beta_i c_1 - \alpha_i c_2)(\gamma c_2 - \beta_i c_3)}$ *and* $w_i = \frac{\alpha_i \beta_i c_3^2}{(\gamma c_1 - \alpha_i c_3)(\gamma c_2 - \beta_i c_3)}$.

Proof. We ignore the proof of Corollary 4.2 since it can be obtained through simple integral calculations. □

Corollary 4.2 gives the probability distribution functions of the service time and complementary service time distributions. Next, we will use the relevant results of the $M/G/1$ model to explore the mean waiting time of residents in different queues. Let $E[W_h]$ and $E[W_l]$ be the mean waiting times of residents in the HQ and LQ, respectively. The following theorem gives the results for the mean waiting time $E[W_h]$ and $E[W_l]$.

Theorem 4.3. *In the vaccine ordering queueing system, the mean waiting time for residents in the HQ is*

$$E[W_h] = \frac{1 - P_{01} \sum\limits_{j=0}^{\infty} \prod\limits_{m=0}^{j} \theta\lambda \int_0^{\infty} \mathrm{e}^{-(m+1)i_h x} \bar{F}_1(x)\mathrm{d}x - P_{01}}{i_h}, \tag{4.4}$$

and the mean waiting time for residents in the LQ is:

$$E[W_l] = \frac{1 - P_{02} \sum\limits_{j=0}^{\infty} \prod\limits_{m=0}^{j} (1-\theta)\lambda \int_0^{\infty} \mathrm{e}^{-(m+1)i_l x} \bar{F}_2(x)\mathrm{d}x - P_{02}}{i_l}, \tag{4.5}$$

where

$$P_{01} = \frac{1}{\sum\limits_{j=0}^{\infty} \prod\limits_{m=0}^{j} \theta\lambda \int_0^{\infty} \mathrm{e}^{-m i_h x} \bar{F}_1(x)\mathrm{d}x + 1}, \tag{4.6}$$

and

$$P_{02} = \frac{1}{\sum\limits_{j=0}^{\infty} \prod\limits_{m=0}^{j} (1-\theta)\lambda \int_0^{\infty} \mathrm{e}^{-m i_l x} \bar{F}_2(x)\mathrm{d}x + 1}. \tag{4.7}$$

Proof. In the higher-efficacy vaccine queue, the probability density function of virtual waiting time is:

$$f_h(x) = \theta\lambda \bar{F}_1(x) P_{01} + \theta\lambda \int_0^x \bar{F}(x-y) \mathrm{e}^{-i_h y} f_h(y) \mathrm{d}y, \tag{4.8}$$

where $P_{01} + \int_0^\infty f_h(x)\mathrm{d}x = 1$. The Laplace transform of (4.8) is:

$$\tilde{f}_h(s+i_h) - \frac{\tilde{f}_h(s)}{\theta\lambda\tilde{\beta}_1(s)} = -P_{01},$$

where $\tilde{f}_h(s+i_h) = \int_0^\infty \mathrm{e}^{-(s+i_h)x} f_h(x)\mathrm{d}x$, $\tilde{f}_h(s) = \int_0^\infty \mathrm{e}^{-sx} f_h(x)\mathrm{d}x$ and $\tilde{\beta}_1(s) = \int_0^\infty \mathrm{e}^{-sx}\bar{F}_1(x)\mathrm{d}x$. Then, we can obtain

$$\tilde{f}_h(s) = P_{01} \sum_{j=0}^{\infty} \prod_{m=0}^{j} \theta\lambda\tilde{\beta}_1(s+mi_h). \tag{4.9}$$

Substituting $s = 0$ into (4.9) and $\tilde{f}_h(s) = \int_0^\infty \mathrm{e}^{-sx} f_h(x)\mathrm{d}x$, we get

$$\tilde{f}_h(0) = P_{01} \sum_{j=0}^{\infty} \prod_{m=0}^{j} \theta\lambda\tilde{\beta}_1(mi_h),$$

and

$$\tilde{f}_h(0) = \int_0^\infty f_h(x)\mathrm{d}x = 1 - P_{01}.$$

Thus, we can derive

$$P_{01} = \frac{1}{\sum\limits_{j=0}^{\infty} \prod\limits_{m=0}^{j} \theta\lambda \int_0^\infty \mathrm{e}^{-mi_h x}\bar{F}_1(x)\mathrm{d}x + 1}. \tag{4.10}$$

From the work of Sarhangian and Balcioglu [33], we have

$$E[W_h] = \frac{1 - \tilde{f}_h(i_h) - P_{01}}{i_h} = \frac{1 - P_{01} \sum\limits_{j=0}^{\infty} \prod\limits_{m=0}^{j} \theta\lambda \int_0^\infty \mathrm{e}^{-(m+1)i_h x}\bar{F}_1(x)\mathrm{d}x - P_{01}}{i_h}. \tag{4.11}$$

Hence, (4.4) and (4.6) have been proved. We ignore the proofs of (4.5) and (4.7) since they can be proved by a similar method. This completes the proof. □

Since residents may be infected during the waiting process, the longer the waiting time of residents, the higher the infection rate of residents during the waiting period. The marginal infection rate is increasing in the mean waiting time. Therefore, an increasing convex function can be used to characterize the relation between the infection rate and the mean waiting time. In real-life situations, it often is verified that the infection rate of residents during the waiting period is exponentially increasing with respect to the mean waiting time. In this paper, we assume that $i_h = \mathrm{e}^{hE[W_h]}$ and $i_l = \mathrm{e}^{hE[W_l]}$, where h is the infection intensity parameter which measures the sensitivity of the infection rate to the waiting time.

Corollary 4.4. *In the vaccine ordering queueing system, the infection rates i_h and i_l of residents during the waiting period are*

$$i_h = exp\left\{ h\frac{\sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}\theta\lambda\int_0^{\infty}\mathrm{e}^{-mi_hx}\bar{F}_1(x)\mathrm{d}x - \sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}\theta\lambda\int_0^{\infty}\mathrm{e}^{-(m+1)i_hx}\bar{F}_1(x)\mathrm{d}x}{i_h\left(\sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}\theta\lambda\int_0^{\infty}\mathrm{e}^{-mi_hx}\bar{F}_1(x)\mathrm{d}x+1\right)} \right\}, \tag{4.12}$$

and

$$i_l = exp\left\{ h\frac{\sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}(1-\theta)\lambda\int_0^{\infty}\mathrm{e}^{-mi_lx}\bar{F}_2(x)\mathrm{d}x - \sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}(1-\theta)\lambda\int_0^{\infty}\mathrm{e}^{-(m+1)i_lx}\bar{F}_2(x)\mathrm{d}x}{i_l\left(\sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}(1-\theta)\lambda\int_0^{\infty}\mathrm{e}^{-mi_lx}\bar{F}_2(x)\mathrm{d}x+1\right)} \right\}. \tag{4.13}$$

Proof. In the HQ and LQ, the infection rates i_h and i_l are relevant to the mean waiting times $E[W_h]$ and $E[W_l]$, and satisfy the following equations:

$$i_h = \mathrm{e}^{hE[W_h]}, \tag{4.14}$$

$$i_l = \mathrm{e}^{hE[W_l]}. \tag{4.15}$$

Substituting (4.4) and (4.6) into (4.14), we can derive (4.12). Similarly, substituting (4.5) and (4.7) into (4.15), we can derive (4.13). This completes the proof. □

Corollary 4.4 presents the relation between the infection rate and the mean waiting time during the waiting process, where $\bar{F}_i$ can be obtained from Corollary 4.2. After a simplified calculation, we can obtain the following results.

Corollary 4.5. *The infection rates i_h and i_l of residents during the waiting period (i.e., (4.12) and (4.13)) can be expressed in the form of a series, namely*

$$i_h = exp\left\{ h\frac{\sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}K_m - \sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}K_{m+1}}{i_h\left(\sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}K_m+1\right)} \right\}, \tag{4.16}$$

and

$$i_l = exp\left\{ h\frac{\sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}L_m - \sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}L_{m+1}}{i_h\left(\sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}L_m+1\right)} \right\}, \tag{4.17}$$

where

$$K_m = u_1\theta\lambda\frac{1}{mi_h+\frac{\alpha_1}{c_1}} - v_1\theta\lambda\frac{1}{mi_h+\frac{\beta_1}{c_2}} + w_1\theta\lambda\frac{1}{mi_h+\frac{\gamma}{c_3}}, \tag{4.18}$$

$$L_m = u_2(1-\theta)\lambda\frac{1}{mi_l+\frac{\alpha_2}{c_1}} - v_2(1-\theta)\lambda\frac{1}{mi_l+\frac{\beta_2}{c_2}} + w_2(1-\theta)\lambda\frac{1}{mi_l+\frac{\gamma}{c_3}}, \tag{4.19}$$

and $u_i = \frac{\beta_i\gamma c_1^2}{(\beta_i c_1-\alpha_i c_2)(\gamma c_1-\alpha_i c_3)}$, $v_i = \frac{\alpha_i\gamma c_2^2}{(\beta_i c_1-\alpha_i c_2)(\gamma c_2-\beta_i c_3)}$, $w_i = \frac{\alpha_i\beta_i c_3^2}{(\gamma c_1-\alpha_i c_3)(\gamma c_2-\beta_i c_3)}$, *(i=1,2).*

Proof. According to (4.3), we can compute $\int_0^\infty \mathrm{e}^{-mi_h x}\bar{F}_1(x)\mathrm{d}x$, *i.e.*,

$$\begin{aligned}\int_0^\infty \mathrm{e}^{-mi_h x}\bar{F}_1(x)\mathrm{d}x &= \int_0^\infty \mathrm{e}^{-mi_h x}(u_1\mathrm{e}^{-\frac{\alpha_1}{c_1}x} - v_1\mathrm{e}^{-\frac{\beta_1}{c_2}x} + w_1\mathrm{e}^{-\frac{\gamma}{c_3}x})\mathrm{d}x,\\ &= u_1\frac{1}{mi_h+\frac{\alpha_1}{c_1}} - v_1\frac{1}{mi_h+\frac{\beta_1}{c_2}} + w_1\frac{1}{mi_h+\frac{\gamma}{c_3}}.\end{aligned} \tag{4.20}$$

Furthermore, we can calculate $\int_0^\infty \mathrm{e}^{-(m+1)i_h x}\bar{F}_1(x)\mathrm{d}x$

$$\begin{aligned}\int_0^\infty \mathrm{e}^{-(m+1)i_h x}\bar{F}_1(x)\mathrm{d}x &= \int_0^\infty \mathrm{e}^{-(m+1)i_h x}(u_1\mathrm{e}^{-\frac{\alpha_1}{c_1}x} - v_1\mathrm{e}^{-\frac{\beta_1}{c_2}x} + w_1\mathrm{e}^{-\frac{\gamma}{c_3}x})\mathrm{d}x,\\ &= u_1\frac{1}{(m+1)i_h+\frac{\alpha_1}{c_1}} - v_1\frac{1}{(m+1)i_h+\frac{\beta_1}{c_2}} + w_1\frac{1}{(m+1)i_h+\frac{\gamma}{c_3}}.\end{aligned} \tag{4.21}$$

Similarly, we can calculate $\int_0^\infty \mathrm{e}^{-mi_h x}\bar{F}_2(x)\mathrm{d}x$ and $\int_0^\infty \mathrm{e}^{-(m+1)i_h x}\bar{F}_2(x)\mathrm{d}x$. Then, substituting (4.20)-(4.21) into (4.12), we can obtain

$$i_h = exp\left\{h\frac{\sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}K_m - \sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}K_{m+1}}{i_h\left(\sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}K_m+1\right)}\right\}.$$

Using a similar method, we get

$$i_l = exp\left\{h\frac{\sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}L_m - \sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}L_{m+1}}{i_h\left(\sum\limits_{j=0}^{\infty}\prod\limits_{m=0}^{j}L_m+1\right)}\right\}.$$

This completes the proof of Corollary 4.5. □

Corollary 4.5 gives the relation between i_h and K_m (or i_l and L_m), where K_m (or L_m) is determined by θ and i_h (or i_l). (4.16) and (4.17) are expressions shaped as $x = f(x)$, where x is a fixed point. In this paper, we perform a simple iteration to obtain i_h (or i_l) by using the Matlab software. In Theorem 4.3 and Corollary 4.4, we can get the performance measures of the system and the infection rates of residents waiting in HQ and LQ.

Based on the above results, the total infection rate, denoted by I_a, can be derived as follows:

$$I_a = \theta \left[i_h + \frac{(\lambda\theta - i_h)i_{hf}}{\lambda\theta} \right] + (1-\theta) \left[i_l + \frac{(\lambda(1-\theta) - i_l)i_{lf}}{\lambda(1-\theta)} \right], \tag{4.22}$$

which will help to explore the socially optimal ordering strategy.

4.2. Socially optimal ordering strategy

Under the service capacity regime, the service capacity μ_1 and μ_2 are fixed. At this time, the total infection rate can be obtained from (4.22), where both the infection rates i_h and i_l in the waiting process are functions of the vaccine ordering policy θ, then I_a is dependent of θ. For the convenience of analysis, I will mark I_a as $I_a(\theta)$. In order to find the socially optimal ordering strategy θ^* that minimizes the total infection rate, we need to solve the following optimization problem:

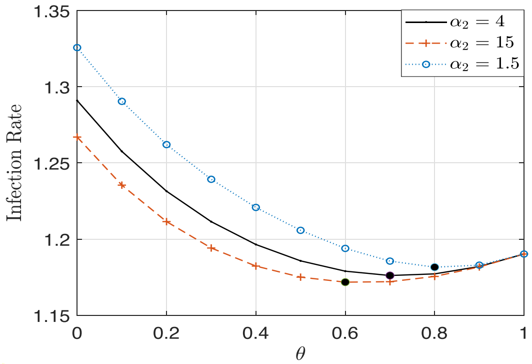
$$\begin{cases} \arg\min I_a[\theta] = \theta \left[i_h[\theta] + \dfrac{(\lambda\theta - i_h[\theta])i_{hf}}{\lambda\theta} \right] + (1-\theta) \left[i_l[\theta] + \dfrac{(\lambda(1-\theta) - i_l[\theta])i_{lf}}{\lambda(1-\theta)} \right], \\ s.t.\ 0 \le \theta \le 1. \end{cases} \tag{4.23}$$

The above optimization problem has crucial implications for managers attempting to effectively prevent the spread of the COVID-19 virus. In the following, we find the socially optimal ordering strategy θ^* for each parameter taking different values by numerical experiments. We consider two cases, *i.e.*, the cases of the small and large arrival rates. Based on these two cases, we analyze the trend of each parameter θ^* and the minimum total infection rate I_a^* ($I_a^* = min(I_a)$), which is dependent of i_h and i_l. Due to the complexity of the calculation of parameters i_h and i_l, we use a Matlab program to obtain the corresponding numerical results.

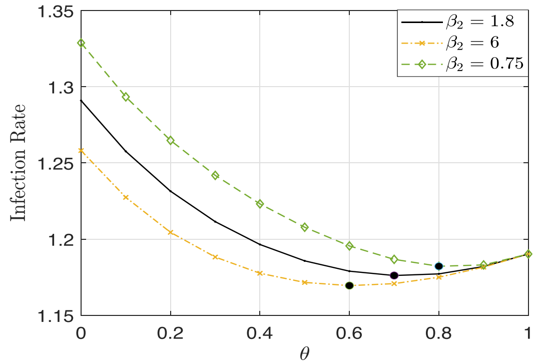
• **The case of small resident arrival rate**. By observing Figure 3a-d, we can find that the total infection rate I_a is convex concerning θ as $\lambda = 5$. Therefore there exists an optimal vaccine ordering strategy θ^* that minimizes the total social infection rate I_a. For example, the socially optimal ordering strategy is $\theta^* = 0.7$ for λ=5, α_1=2, β_1=0.9, α_2=4, β_2=1.8, γ=1.4, i_{hf}=0.05, i_{lf}=0.2, h=0.3. The optimal ordering strategy is usually at a certain value within the interval of $(0, 1)$ (see Fig. 3a–c), which indicates that in most cases we should adopt a mixed ordering strategy, that is, ordering a certain proportion of higher-efficacy vaccine and the remaining proportion of lower-efficacy vaccine. However, under certain extreme conditions, ordering only a certain type of vaccine is an optimal strategy for social managers. For example, when i_{lf} is large (see the case of i_{lf}=0.8 in Fig. 3d), the socially optimal ordering strategy may be to order higher-efficacy vaccine. This is also quite intuitive, because from the perspective of social managers, if the infection rate among residents who have completed the lower-efficacy vaccination is still high, they would rather have residents spend longer waiting times than receive the lower-efficacy vaccine. Figure 3d shows that the total infection rate I_a decreases monotonically with θ when $i_{lf} = 0.8$. In this situation, the socially optimal ordering strategy θ^*=1, *i.e.*, the socially optimal vaccine ordering strategy is only to order the higher-efficacy vaccine.

• **The case of large resident arrival rate**. When the resident arrival rate is very high, the overall infection rate will no longer be concave, and there are multiple extreme points in the situation. When exploring the optimal ordering strategy, it is necessary to compare the function values corresponding to these extreme points. Figure 4a–d shows the function curve of the total infection rate as λ=30, and there may exist one or more minimum points for the total infection rate. For example, there are two minimum points for λ=30, α_1=2, β_1=0.9, α_2=4, β_2=1.8, γ=1.4, i_{hf}=0.05, i_{lf}=0.2, h=0.3 (see Fig. 4a) and the socially optimal ordering strategy θ^* is 0.75, while there is only one minimum point for λ=5, α_1=2, α_2=4, β_1=0.9, β_2=1.8, γ=1.4, i_{hf}=0.05, i_{lf}=0.8, h=0.3 (see Fig. 4d) and the socially optimal ordering strategy $\theta^* = 1$.

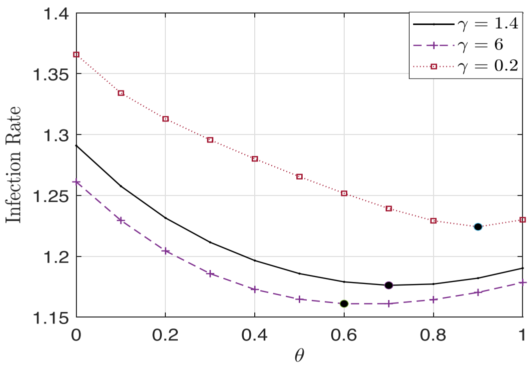
The impact of different parameters on the socially optimal ordering strategy is also a practical issue worth exploring. To explore the effect of parameters λ, α_2, β_2, γ, i_{hf} and i_{lf} on θ^*, we offer an essential numerical experiment (see Fig. 5). According to Figures 5a and 5f, we can find that the optimal parameter θ^* increases



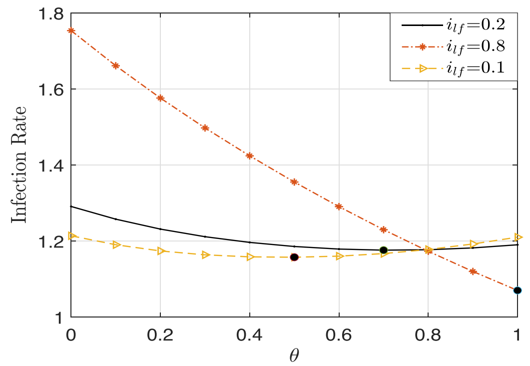
(a) α_1=2, β_1=0.9, β_2=1.8, γ=1.4, i_{hf}=0.05, i_{lf}=0.2, h=0.3



(b) α_1=2, α_2=4, β_1=0.9, γ=1.4, i_{hf}=0.05, i_{lf}=0.2, h=0.3



(c) α_1=2, α_2=4, β_1=0.9, β_2=1.8, i_{hf}=0.05, i_{lf}=0.2, h=0.3



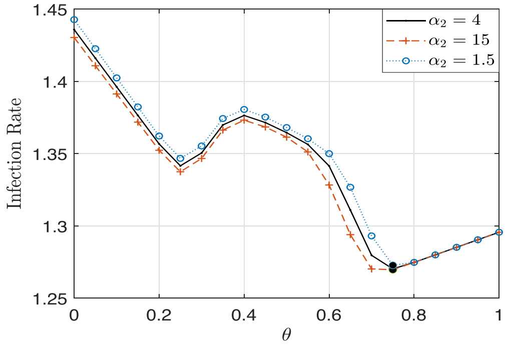
(d) α_1=2, α_2=4, β_1=0.9, β_2=1.8, γ=1.4, i_{hf}=0.05, h=0.3

FIGURE 3. Infection rate I_a vs. θ as λ=5.

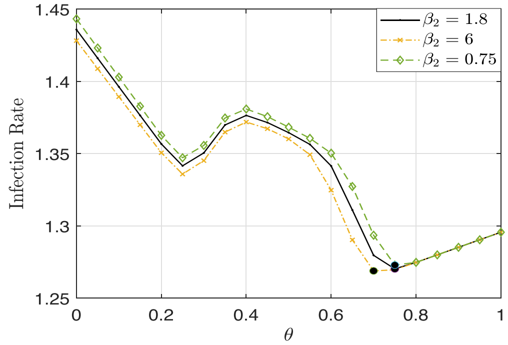
with λ and i_{lf} , while θ^* decreases with parameters α_2, β_2, γ and i_{hf} (see Fig. 5b–e). For instance, as the arrival rate of residents grows (Fig. 5a), the system becomes more crowded and residents are at increased risk of infection during the waiting process. In this situation, a mixed vaccine order strategy is better, which can help alleviate congestion. However, when the arrival rate is large enough that the risk of infection is high in both queues, it is better to make residents receive higher quality services, *i.e.*, residents are vaccinated with efficient vaccines. Similarly, when the production capacity α_2 of lower-efficacy vaccine increases (Fig. 5b), the LQ becomes more serviceable and residents are able to get vaccinated faster in the LQ. In this scenario, the optimal vaccine ordering strategy θ^* decreases, *i.e.*, more low-efficacy vaccines are ordered.

5. INVESTMENT COST REGIME

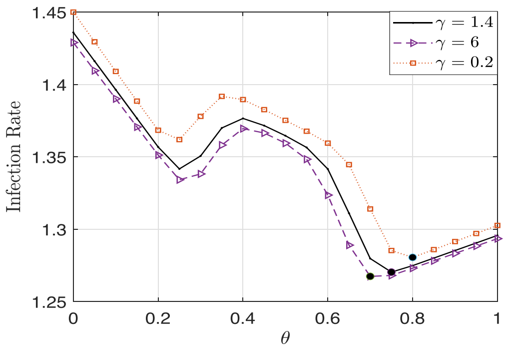
In reality, the government budget is a significant aspect of vaccine ordering system management since it establishes the upper limit of service capacity (efficiency) and service quality. After determining the ordering strategy of the higher- and lower-efficacy vaccines, the service rates of HQ and LQ can be enhanced by a specific expense to control the spread of the COVID-19 virus as soon as possible. If the speed of service is faster in the HQ and LQ, the shorter the waiting time of residents, the less the probability that they will become infected while waiting. However, governments cannot have unlimited assets to speed up services because their budgets are restricted. As a result, we assume in this section that the government has a budget of Q per unit of time to improve service capacity. As introduced earlier, the service capacity of higher-efficacy/lower-efficacy



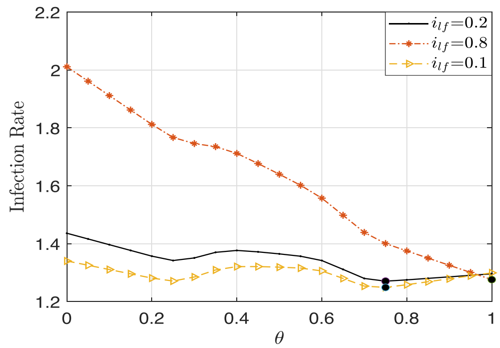
(a) α_1=2, β_1=0.9, β_2=1.8, γ=1.4, i_{hf}=0.05, i_{lf}=0.2, h=0.3



(b) α_1=2, α_2=4, β_1=0.9, γ=1.4, i_{hf}=0.05, i_{lf}=0.2, h=0.3



(c) α_1=2, α_2=4, β_1=0.9, β_2=1.8, i_{hf}=0.05, i_{lf}=0.2, h=0.3



(d) α_1=2, α_2=4, β_1=0.9, β_2=1.8, γ=1.4, i_{hf}=0.05, h=0.3

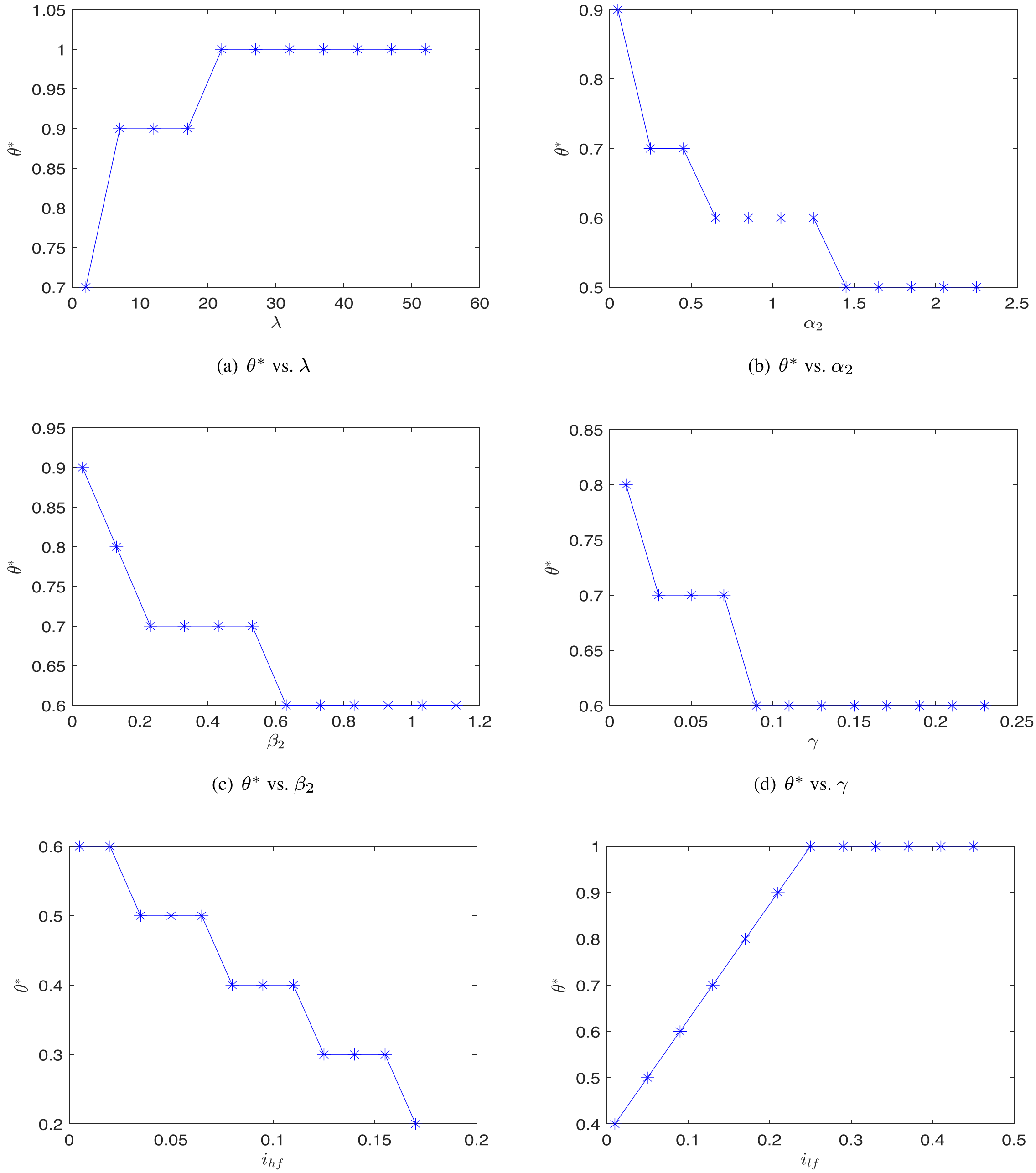
FIGURE 4. Infection rate I_a vs. θ as λ=30.

vaccines is influenced by three variables: the capacity to produce vaccines, the capacity to transport vaccines, and the community service capacity. Therefore, improving the service capacity of this system can be achieved by increasing the production and transportation capacity of the vaccines and the community service capacity, *i.e.*, by optimizing α_1, β_1, α_2, β_2 and γ at a certain cost.

Our overall goal remains to minimize the total infection rate. However, unlike Section 4, in this section we reduce residents' waiting times by spending a fixed amount to improve the service capacity of HQ and LQ. By allowing residents to get vaccinated more quickly, the overall infection rate in the community is reduced. In the context of a COVID-19 pandemic, vaccine demand exceeds supply at a time when the capacity to produce vaccines may already be at its maximum, so we assume that the production capacity is at its maximum and that increasing costs have no effect on the capacity. Then, increasing service rates can be achieved by increasing the capacity to transport vaccines β_1, β_2 and the capacity of community services γ. We assume that c_{h1}, c_{h2}, c_{l1}, c_{l2}, c_s mean the cost per unit time to increase the production capacity of the vaccine in the HQ, the cost to transport the vaccine in the HQ, the cost to produce the vaccine in the LQ, the cost to transport the vaccine in the LQ, and the cost to serve the community in the HQ and LQ, respectively.

Under the condition that the total budget and the ordering strategy are fixed, the total infection rate can be obtained by (4.22), which is a function of β_1, β_2 and γ. Thus there is the following optimization problem:

$$\begin{cases} \mathrm{argmin} I_a[\beta_1,\beta_2,\gamma] = \dfrac{(i_h[\beta_1,\beta_2,\gamma](\lambda\theta - i_{hf}) + \lambda\theta i_{hf}) + (i_l[\beta_1,\beta_2,\gamma](\lambda(1-\theta) - i_{lf}) + \lambda(1-\theta) i_{lf})}{\lambda}, \\ s.t.\ Q = c_{h1}\alpha_1 + c_{h2}\beta_1 + c_{l1}\alpha_2 + c_{l2}\beta_2 + c_s\gamma. \end{cases}$$



(a) θ^* vs. λ

(b) θ^* vs. α_2

(c) θ^* vs. β_2

(d) θ^* vs. γ

(e) θ^* vs. i_{hf}

(f) θ^* vs. i_{lf}

FIGURE 5. The effect of parameters λ, α_2, β_2, γ, i_{hf} and i_{lf} on θ^* for λ=0.5, α_1=0.5, β_1=0.45, α_2=0.75, β_2=1.2, γ=0.6, i_{hf}=0.005, i_{lf}=0.1, h=0.3.

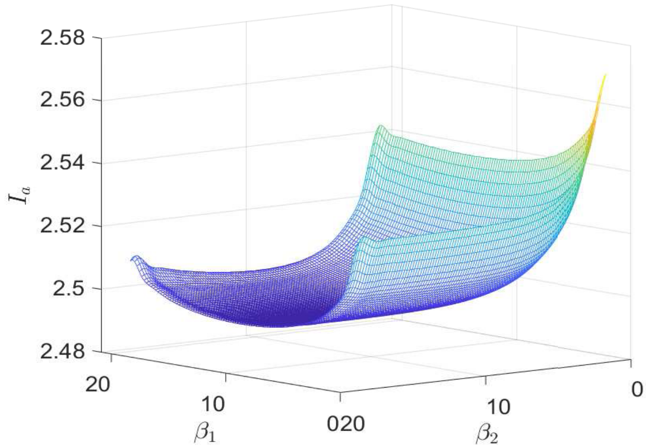


FIGURE 6. Infection rate vs. β_1 and β_2, for λ=50, α_1=2, α_2=4, Q=20, c_{h1}=1, c_{h2}=0.5, c_{l1}=0.3, c_{l2}=0.2, c_s=1, i_{hf}=0.05, i_{lf}=0.2, h=3, θ=0.8.

TABLE 2. Infection rate vs. β_1 and β_2, for λ=50, α_1=2, α_2=4, Q=20, c_{h1}=1, c_{h2}=0.5, c_{l1}=0.3, c_{l2}=0.2, c_s=1, i_{hf}=0.05, i_{lf}=0.2, h=3, θ=0.8.

I_a \ β_2 / β_1	10	11	12	13	14	15
0.9	2.526855958	2.526703604	2.526596838	2.52652062	2.526465636	2.526425904
1.9	2.512866635	2.512722274	2.512624983	2.512559436	2.512516126	2.512488948
2.9	2.505870059	2.505730437	2.505639396	2.505581297	2.505546429	2.505528551
3.9	2.501747347	2.50161056	2.501523937	2.501471506	2.501443337	2.501433041
4.9	2.499086835	2.498951829	2.498868680	2.498821062	2.498798809	2.498795376
5.9	2.497275315	2.497141534	2.497061443	2.497018341	2.497001812	2.497005145
6.9	2.496004847	2.495872064	2.495794973	2.495756480	2.495745902	2.495756354
7.9	2.495105153	2.494973395	2.494899539	2.494866070	2.494862026	2.494880343
8.9	2.494475777	2.494345312	2.494275201	2.494247490	2.494250927	2.494278259
9.9	2.494055075	2.493926435	2.493860893	2.493840037	2.493852314	2.493890284
10.9	2.493804776	2.493678814	2.493619048	2.493606594	2.493629595	2.493680425
11.9	2.493701826	2.493579827	2.493527554	**2.493525648**	2.493561949	2.493628670
12.9	2.493734008	2.493617853	2.493575496	2.493587109	2.493640257	2.493727021
13.9	2.493897693	2.493790117	2.493761104	2.493790394	2.493865319	2.493977902

Figure 6 shows how the overall infection rate function I_a varies with β_1 and β_2 for a given set of parameters. It can be observed that I_a is a convex function with respect to β_1 and β_2, so there is (β_1^*, β_2^*) such that the total infection rate I_a is minimized. From Table 2 we can find that the total infection rate I_a is minimized when β_1=11.9 and β_2=13. Therefore, the optimal transport rates of vaccines β_1^*=11.9, β_2^*=13. From the constraint condition of $Q = c_{h1}\alpha_1 + c_{h2}\beta_1 + c_{l1}\alpha_2 + c_{l2}\beta_2 + c_s\gamma$, the corresponding optimal community service rate γ^* can be obtained, *i.e.*, γ^*=8.25.

6. Conclusions and future works

In the context of COVID-19, this paper discusses the optimal vaccine ordering strategy under the service capability regime and the optimization of service capability under the investment cost regime. Under the service capacity regime, the service rates of HQ and LQ are determined, and social managers achieve the lowest total infection rate in society by selecting appropriate vaccine ordering strategies. However, under the investment cost regime, the total investment cost of social managers is fixed, and they achieve the lowest total social infection rate through a reasonable allocation of investment costs.

Our model is not only applicable to the ordering strategy of the COVID-19 vaccine and the investment decision-making of social managers, but it can also help the decision-making of other infectious diseases. In future work, we will explore the following aspects.

• **Strategic choice of vaccination for residents**. In this paper, we did not consider the residents' strategic selection of higher- and lower-efficacy vaccines, since this paper is based on the COVID-19 background to discuss the decision-making problem of social managers. During the epidemic period of COVID-19, the vaccine shortage is quite serious, and it is difficult for residents to make a choice based on their own preferences for different vaccines. However, in other infectious disease vaccination issues, residents may have multiple choices, and they may choose vaccines with different protection efficiency. In this scenario, the behavioral strategies of residents are worth studying. A fitting example is the choice of human papillomavirus (HPV) vaccine. There are bivalent vaccines (Cervarix), quadrivalent vaccines (Gardasil-4), and a nine-valent vaccine (Gardasil 9) for preventing cervical cancer. Morris [34] suggests selecting vaccines based on the epidemiological characteristics of different regions. However, residents have the right to choose whether to receive the HPV vaccine and the type of vaccine, and their specific action strategies depend on their level of risk preference. To the best of our knowledge, no scholar has used queueing theory to study the behavioral decisions of risk sensitive residents receiving vaccines.

• **Optimal strategies for different stakeholders in multi-party games**. Based on the fact that vaccine manufacturers practically had a monopoly in the early and middle stages of the COVID-19 outbreak, this paper did not analyze vaccine manufacturers' investment decisions. However, in the case of other infectious disease immunizations, the game among residents, service providers, and social managers must be investigated. In real-life situations, when an epidemic breaks out, residents will weigh vaccine prices and effectiveness to make choices that maximize individual benefits. As for the companies responsible for production, transportation, etc., their goal is to maximize corporate profits, which will affect the price of vaccines. Furthermore, for social managers, maximizing social benefits is their ultimate goal. In this situation, the ideal strategic decision among the three parties is an interesting practical issue. Li *et al.* [35] discussed a similar tripartite game problem, and based on the background of carbon trading policy, they studied a tripartite evolutionary game model among developers, governments, and consumers in the green building market. In the future, we will study the tripartite game among residents with the right to choose, companies producing and transporting vaccines, and social managers to obtain a "win-win-win" strategy that optimizes all three parties.

Acknowledgements. This research was supported in part by the National Natural Science Foundation of China (Grant Nos. 72371259, 71871008), the Emerging Interdisciplinary Project of CUFE (21XXJC010), the Emerging Interdisciplinary Project of CUFE (21XXJC010), the Project of Establishing the "Double First-Class" Discipline of Surveying and Mapping Science and Technology (Grant No. GCCRC202307), and the Fundamental Research Funds for the Universities of Henan Province (Grant No. NSFRF220442).

References

[1] J.C. Castillo, A. Ahuja, S. Athey, *et al.*, Market design to accelerate COVID-19 vaccine supply. *Science* **371** (2021) 1107–1109.
[2] M. Wadman, The overlooked superpower of mRNA vaccines. *Science* **373** (2021) 479–479.
[3] M.N. Uddin and M.A. Roni, Challenges of storage and stability of mRNA-based COVID-19 vaccines. *Vaccines* **9** (2021) 1033.
[4] J. Cohen, What went wrong with CureVac's mRNA vaccine? *Science* **372** (2021) 1381–1381.
[5] C. Huang, Y. Wang, X. Li, *et al.*, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395** (2020) 497–506.
[6] D.S.W. Ting, L. Carin, V. Dzau and T.Y. Wong, Digital technology and COVID-19. *Nat. Med.* **26** (2020) 459–461.
[7] M. Rozhkov, D. Ivanov, J. Blackhurst and A. Nair, Adapting supply chain operations in anticipation of and during the COVID-19 pandemic. *Omega* **110** (2022) 1–17.
[8] S.M. Hosseini-Motlagh, M.R.G. Samani and S. Homaei, Design of control strategies to help prevent the spread of COVID-19 pandemic. *Eur. J. Oper. Res.* **304** (2023) 219–238.
[9] Y. Perlman and U. Yechiali, Reducing risk of infection – the COVID-19 queueing game. *Saf. Sci.* **132** (2020) 1–13.

[10] L. Yang, S. Cui and Z. Wang, Design of COVID-19 testing queues. *Prod. Oper. Manag.* **31** (2022) 2204–2221.

[11] K. Kang, S. Doroudi, M. Delasay and A. Wickeham, A queueing-theoretic framework for evaluating transmission risks in service facilities during a pandemic. *Prod. Oper. Manag.* **32** (2023) 1453–1470.

[12] L. Nageswaran, Implications of vaccine shopping during pandemic. *Prod. Oper. Manag.* **32** (2023) 1133–1149.

[13] L. Tang, Y. Li, D. Bai, T. Liu and L.C. Coelho, Bi-objective optimization for a multi-period COVID-19 vaccination planning problem. *Omega* **110** (2022) 1–18.

[14] S.E. Martonosi, B. Behzad and K. Cummings, Pricing the COVID-19 vaccine: a mathematical approach. *Omega* **103** (2021) 1–7.

[15] K. Cummings, B. Behzad and S. Martonosi, Centers for disease control and prevention as a strategic agent in the pediatric vaccine market: an analytical approach. *Manuf. Serv. Oper. Manag.* **23** (2021) 1398–1412.

[16] D.Y. Barrer, Queuing with impatient customers and indifferent clerks. *Oper. Res.* **5** (1957) 644-649.

[17] D.Y. Barrer, Queuing with impatient customers and ordered service. *Oper. Res.* **5** (1957) 650–656.

[18] J.C.J. Ancker and A.V. Gafarian, Some queuing problems with balking and reneging. I. *Oper. Res.* **11** (1963) 88–100.

[19] V. Sarhangian and B. Balcioglu, Waiting time analysis of multi-class queues with impatient customers. *Probab. Eng. Inform. Sci.* **27** (2013) 333–352.

[20] A. Economou, D. Logothetis and A. Manou, The value of reneging for strategic customers in queueing systems with server vacations/failures. *Eur. J. Oper. Res.* **299** (2022) 960–976.

[21] C.H. Tuohy, C.M. Flood and M. Stabile, How does private finance affect public health care systems? Marshaling the evidence from OECD nations. *J. Health. Polit. Polic.* **29** (2004) 359–396.

[22] Z.G. Zhang, Performance analysis of a queue with congestion-based staffing policy. *Manage. Sci.* **55** (2009) 240–251.

[23] S. Zhu, J. Wang and W.W. Li, Cloud or in-house service? strategic joining and social optimality in hybrid service systems with retrial orbit. *IEEE. Syst. J.* **17** (2023) 3810–3821.

[24] P. Guo, R. Lindsey and Z.G. Zhang, On the Downs–Thomson paradox in a self-financing two-tier queuing system. *Manuf. Serv. Oper. Manag.* **16** (2014) 315–322.

[25] Z. Hua, W. Chen and Z.G. Zhang, Competition and coordination in two-tier public service systems under government fiscal policy. *Prod. Oper. Manag.* **25** (2016) 1430–1448.

[26] J. Wang, Z. Wang, Z.G. Zhang and F. Wang, Efficiency-quality trade-off in allocating resource to public healthcare systems. *Int. J. Prod. Res.* **60** (2022) 6469–6490.

[27] G. Nicosia, A. Pacifici and U. Pferschy, Price of fairness for allocating a bounded resource. *Eur. J. Oper. Res.* **257** (2017) 933–943.

[28] K. Zhang, X. Hong, Y. Han and B. Wang, Optimal discrete resource allocation on metapopulation networks for suppressing spatial spread of epidemic. *Chaos. Soliton. Fract.* **169** (2023) 1–11.

[29] J. Chen, T. Du and G. Xiao, A multi-objective optimization for resource allocation of emergent demands in cloud computing. *J. Cloud. Comput.* **10** (2021) 1–17.

[30] C. Franco, N. Herazo-Padilla and J. A. Castaneda, A queueing Network approach for capacity planning and patient Scheduling: a case study for the COVID-19 vaccination process in Colombia. *Vaccine* **40** (2022) 7073–7086.

[31] H. Jahani, A.E. Chaleshtori, S.M.S. Khaksar, A. Aghaie and J.B. Sheu, COVID-19 vaccine distribution planning using a congested queuing system-A real case from Australia. *Transport. Res. E-log.* **163** (2022) 1–33.

[32] S.L. Song, Derivation of an independent random variables and distributions subject to exponential distribution. *J. Liaoning. Normal. Univ.* **4** (1990) 51–58.

[33] V. Sarhangian and B. Balcioglu, Waiting time analysis of multi-class queues with impatient customers. *Probab. Eng. Inform. Sci.* **27** (2013) 333–352.

[34] S.R. Morris, HPV vaccine strategies: the cost of HPV and the choice of vaccine. *Sex. Transm. Infect.* **85** (2009) 315–316.

[35] H. Li, D. Liu, Z.J. Han, Y.P. Sun, L. Wang and J.S. Zhang, Simulation analysis of the three-party evolutionary game of green building market players under carbon trading policy. *Environ. Sci. Pollut. Res.* (2023) 1–18.

Please help to maintain this journal in open access!

This journal is currently published in open access under the Subscribe to Open model (S2O). We are thankful to our subscribers and supporters for making it possible to publish this journal in open access in the current year, free of charge for authors and readers.

Check with your library that it subscribes to the journal, or consider making a personal donation to the S2O programme by contacting subscribers@edpsciences.org.

More information, including a list of supporters and financial transparency reports, is available at https://edpsciences.org/en/subscribe-to-open-s2o.