Safety Science 138 (2021) 105194

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

The impact of infection risk on customers' joining strategies

Yael Perlman [a,*], Uri Yechiali [b]

[a] Department of Management, Bar Ilan University, Ramat Gan 52900, Israel
[b] Department of Statistics and Operations Research, School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel

ARTICLE INFO

Keywords:
Queueing
Risk
Strategic decisions
Threshold strategy

ABSTRACT

The risk of infection from the COVID-19 virus dictates businesses, such as supermarkets and department stores, to impose limits on the maximal number of customers allowed inside a store at any given time. These social distancing constraints generate long queues of waiting customers outside such businesses. This work investigates the impact of infection risk on arriving customers' strategic decisions regarding joining such queues. We consider a typical store where the floor is divided into two separate areas: (i) a shopping area with at most K shoppers allowed, and (ii) a payment area with $c \geq 1$ parallel servers and an adjacent limited waiting space of size $N \geq 0$. When the shopping area is full, a newly arriving customer observes only the outside queue and decides whether to join or balk. We investigate customers' individual joining strategies, as well as social optimization, with a utility function that takes into account not only the cost associated with waiting times (as in Naor's (1969) celebrated model), but also the cost related to the risk of infection. We propose an innovative risk measure that is a function of both the number of customers already in line, and those that a tagged customer 'meets' while waiting to enter the store. Consequently, expressions for mean waiting times and infection risk are derived and explicit formulas are obtained for limit values of the parameters. Our results can be used by authorities and customers alike to determine the maximal allowed queue sizes that ensure safety and reduce the risk of infection while minimizing associated costs.

1. Introduction

The COVID-19 pandemic has forced numerous businesses, including department stores and supermarkets, to limit the number of shoppers occupying a store at any given time to minimize the risk of infection. Due to these restrictions, long queues are formed outside these businesses. Clearly, the more people in the queue, the greater is the likelihood of infection as it becomes harder to ensure social distancing (Long et al., 2020). Arriving customers may refrain from joining long queues, basing their decisions in large part on the *risk of infection* induced while waiting in line.

In this research, we construct and analyze a multi-server queueing model involving two interlaced service phases that limits the number of customers in each phase and takes internal blocking and delay into account. Customers arrive randomly at the store and, when finding a queue outside (i.e., the store is occupied by the maximum number of allowed customers), decide whether to balk or to wait outside the store (in a first-come, first-served (FCFS) line) until permitted to enter. After receiving permission to enter, every customer passes through two service phases: shopping and payment. A customer first spends a random period of time in the shopping area and then proceeds to the payment area (second phase) where cashiers are assigned to serve customers. To control waiting times in the payment queue and ensure the safety of cashiers and customers, store management sets aside a separate waiting space with limited capacity as part of the payment area. When the payment area is full (all cashiers are busy and all waiting spaces are occupied), a customer who has completed shopping "orbits" in the shopping area for a random period of time before trying again to enter the payment area (see, e.g., Avrachenkov et al., 2014; Perel and Yechiali, 2014). When the payment area is not full, a customer enters the payment area and becomes a "payer." Then, when at least one cashier is not busy, the payer proceeds directly to a cashier and pays. Otherwise, the payer joins the queue of waiting customers in the payment area.

Though this two-phase service process can appear to be a two-site tandem network (see e.g. Perlman and Yechiali, 2020a), that is not the case. The two service phases are dependent via the imposed upper limits on the total number of customers in each phase. This dependence requires a special probabilistic approach when analyzing the system.

* Corresponding author.
E-mail addresses: yael.perlman@biu.ac.il (Y. Perlman), uriy@tauex.tau.ac.il (U. Yechiali).

https://doi.org/10.1016/j.ssci.2021.105194
Received 12 January 2021; Accepted 28 January 2021
Available online 21 February 2021
0925-7535/

Recently, several queueing models have been constructed to study the impacts of the COVID-19 pandemic (see, e.g., Alban et al., 2020; Kaplan, 2020; Long et al., 2020; Perlman and Yechiali 2020b). The later authors focused on finding the maximal number of customers allowed inside a store, when the risk of infection is proportional to the second factorial moment of the number of customers in the system, $E[L(L-1)]$. While in that model it is assumed that customers do not balk, in the current work the capacity of the store is given and customers may strategically decide whether to join or not. Consequently, our work belongs to the literature on strategic queueing, which has been studied extensively since the pioneering work of Naor (1969), who investigated customers' threshold-type joining strategies in an observable M/M/1 queue. Yechiali (1971, 1972), in an analysis of G/M/1 and G/M/c queues, allowed for state-dependent randomized joining probabilities and showed that optimal non-randomized strategies exist and that, among all non-randomized strategies, a threshold-type joining strategy is optimal. Hassin and Haviv (2003) and Hassin (2016) provide comprehensive surveys of the field. We extend this stream of literature by proposing a novel measure to estimate customers' risk of infection while queueing in crowded areas. The risk is proportional to the number of customers a tagged customer joining the outside queue 'meets' while waiting in line. Based on this risk measure we study strategic behavior of both the customers and of the system's controller (i.e. the authority).

Upon completing shopping, a customer gains a reward but loses opportunity cost due to waiting to be served. While the common assumption in the current literature is that cost is a function of customer's waiting time (or customer's sojourn time), our model considers also customer's cost incurred due to the risk of getting infected. While individual customers are interested in maximizing their own utility, the goal of the authorities is to maximize social welfare - the expected total utility of all the members of the society. I.e., just as customers decide on the maximal queue size beyond which they will not join, the authority decides on the maximum allowed queue size to ensure the overall welfare and minimize the overall risk of infection.

Our contribution is three-fold: (i) constructing and analyzing an unconventional two-site tandem network with inter-dependence between the queues; (ii) proposing and evaluating an infection-risk performance measure that depends on the total number of other customers a waiting customer meets while in line, including future arrivals; and (iii) finding threshold joining strategies of individual customers and of a central authority. These thresholds induce maximum allowed queue sizes from the perspectives of a 'selfish' customer and from the perspective of social welfare.

In contrast to Naor's (1969) model, in our work an individual customer's strategy depends on the actions of other customers. This follows since (i) a customer's waiting time depends on the observable outside queue size, as well as on the number of customers in the shopping and payment areas and (ii) a customer's infection risk depends on the number of customers already in the queue and on the number of customers who will arrive before s/he enters the store. However, we show that the optimal individual threshold value is still higher than the counterpart social-oriented threshold. Insights from our model are useful to authorities and customers alike when making decisions on the maximal allowed queue sizes in order to ensure safety and reduce risk of infection, while minimizing associated costs.

2. Model formulation

Customers arrive at a store according to a Poisson process with rate λ. The shopping process of each customer consists of two stages: shopping and then payment. The *shopping area* has a limited capacity of at most K shopping customers (*shoppers)*. The *payment area* consists of c parallel cashiers and an adjacent area for at most $N \geqslant 0$ waiting customers (*payers*). Payment (service) time of customers is random and is exponentially distributed with mean $1/\mu$. Shopping time is also random and is exponentially distributed with mean $1/\xi$. A customer who arrives when there are K customers already in the shopping area is not allowed to go in and must decide whether to join the queue outside the store and wait (in a FCFS line) until permitted to enter or to balk and cancel the shopping trip.

A shopper who has completed shopping proceeds directly to the payment area. If the number of customers in the payment area is $c+N$, the customer is blocked and must "orbit" in the shopping area for a random period that is distributed exponentially with mean $1/\xi$. If the number of customers in the payment area is less than $c+N$, the customer enters the payment area (entering phase two and becoming a payer). Then, if at least one cashier is idle, the payer proceeds directly to a free cashier and pays. Otherwise, the payer waits in the payment waiting area until a server (cashier) becomes available. All shopping, service, and orbit times are independent of each other and are independent of the arrival process.

As previously stated, this two-phase service process can appear to be a two-site tandem network but is not. The two service stages are dependent via the imposed upper limits on the total number of customers allowed in each phase. Let ${L_1}^{SHOP}$ (${L_1}^{SHOP} = 0,1,2,\dots, K$) denote the number of customers in the shopping area and ${L_1}^{OUT}$ denote the number of customers queueing up outside the store. Let $L_1 = {L_1}^{SHOP} + {L_1}^{OUT}$ and L_2 ($L_2 = 0,1,2,\dots,c+N$) denote the number of customers in the payment area (waiting or being served). Fig. 1 illustrates the system's configuration.

Denote by r the reward gained by a customer who completes shopping and payment. While in Naor's (1969) model, only the cost associated with waiting times were taken into account, we consider an additional cost that is proportional to the risk of COVID-19 infection associated with waiting in queues. The utility U_n of an arriving customer who joins a queue of size $L_1^{OUT} = n$ is given by

$$U_n \equiv r - \gamma W_n - \delta R_n \tag{1}$$

where W_n and R_n denote, respectively, the customer's waiting time and infection risk. γ is the customer's waiting cost per unit of time and δ is a scaling factor associated with the risk of being infected by the virus. It is assumed that utility functions of individual customers are identical and additive from a social point of view.

In Naor's (1969) model, an individual customer's waiting time depended solely on the number of customers ahead in line. In our model, however, a customer's waiting time depends on the number of customers in the shopping area and on the number of customers in the payment area and thus depends on the actions of other customers. Moreover, our measure of the risk of infection depends also on the number of additional customers who arrive during a customer's waiting time outside the store and not only on the number of customers waiting ahead. Thus, an individual customer's strategy depends on the strategy of the other customers. Furthermore, our model deviates from classical, fully observable models in that our arriving customer's strategy is based on observing only the length of the outside queue (i.e., a partially observable two-stage queue).

3. Threshold joining strategies

We analyze threshold-type joining strategies of arriving customers in a partially observable two-stage queue. An arriving customer who observes only the outside queue weighs two alternatives: join the queue or balk. Specifically, if the observed number of customers in the queue outside the store is lower than a threshold, T, the customer joins the outside queue; when the size of the queue is equal to or greater than T, the customer balks. Thus, under such a threshold-type T-strategy, the outside queue is finite with a maximum of T waiting customers. The transition-rate diagram for this system is depicted in Fig. 2.

The system state space is $S = \{(L_1, L_2) = (i,j) : i = 0,1,2,3,\dots,K+T;\ j = 0,1,2,3,\dots,c+N\}$. Let $P_{ij} = P(L_1 = i, L_2 = j)$ denote the steady-state probabilities of this sys-

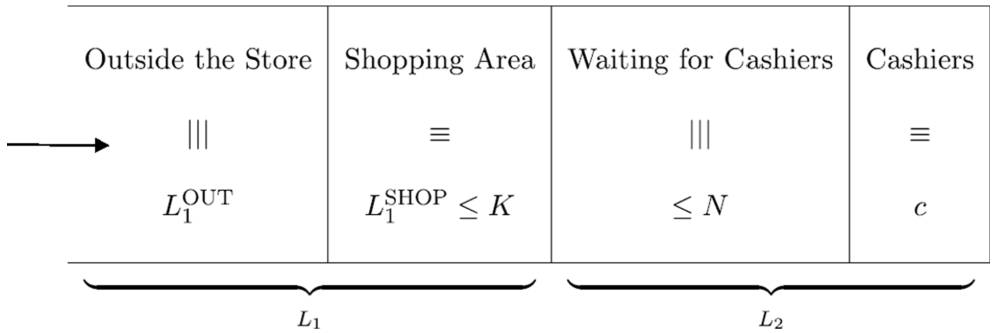

Fig. 1. The two-phase service system configuration.

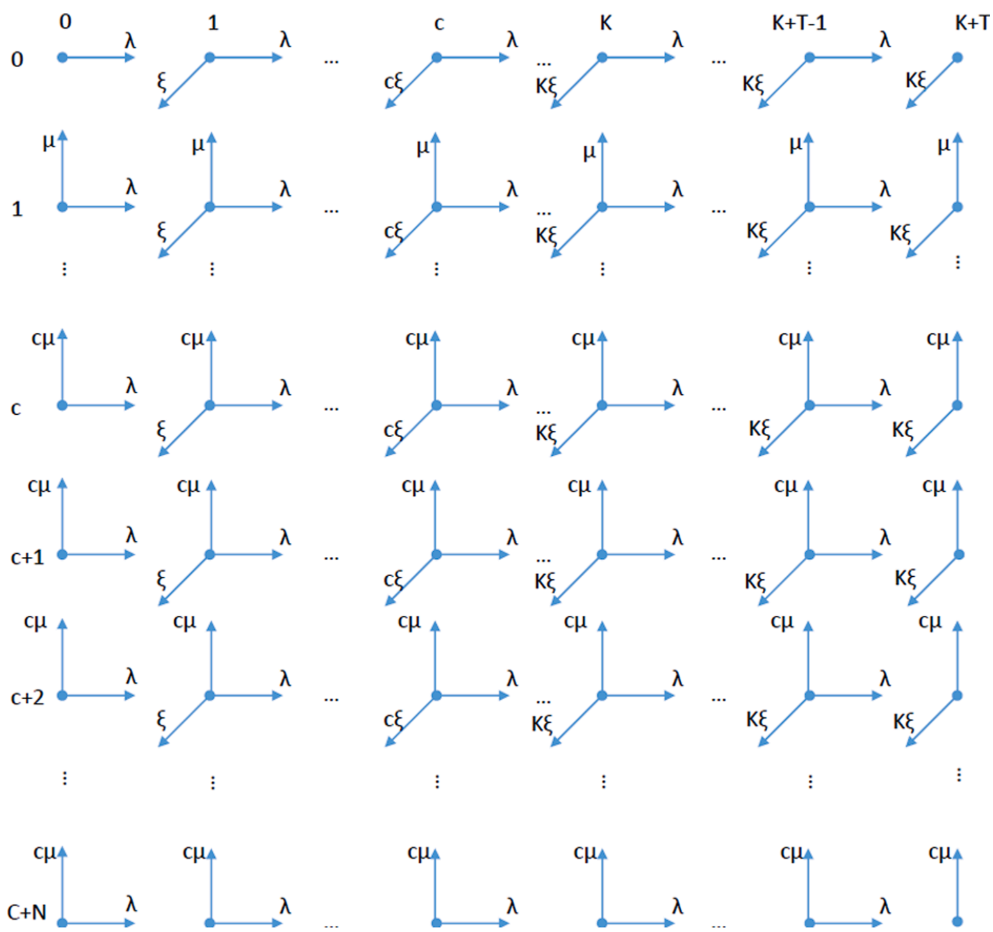

Fig. 2. Transition-rate diagram for the model.

tem's states and let $\overline{P_i} = (P_{i0}, P_{i1}, \ldots, P_{i,c+N})$ denote the probability vector corresponding to the i^{th} column in Fig. 2.

We start by defining two probabilities needed to calculate the expected waiting time of a customer who exercises the T-strategy while all other customers use the same strategy. Consider the case in which the shopping area is full. The system state is $L_1 \geq K$ and $L_2 = j$. Denote by $a(j)$ the probability that a shopper leaves the shopping area before a payer leaves the payment area given that the shopping area is full and there are j customers in the payment area. Then,

$$a(j) = \begin{cases} 0 & j = N + c \\ \dfrac{K\xi}{K\xi + c\mu} & c < j \leqslant N + c - 1 \\ \dfrac{K\xi}{K\xi + j\mu} & 0 \leqslant j \leqslant c \end{cases} \tag{2}$$

Consider a tagged customer who is in position $m+1$ in the outside queue. Define $q_{m,j}$ as the probability that, when the tagged customer becomes the first customer in the outside line, there are no free spaces in the payment area, and j is the number of current payers in the payment area: $j = 0, 1, \ldots, c+N$ (i.e., $L_2 = j$). When $m = 0$, the tagged customer is already the first in line so the number of free spaces in the payment area equals zero only when $L_2 = c + N$. Then,

$$q_{0,j} = \begin{cases} 1 & j = c+N \\ 0 & otherwise \end{cases} \tag{3}$$

For $m \geqslant 1$,

$$q_{m,j} = \begin{cases} 0 & m+j < c+N \\ q_{m-1,j+1} & m \geqslant c+N \text{ and } j=0 \\ a(j)q_{m-1,j+1} + (1-a(j))q_{m,j-1} & m+j \geqslant c+N \text{ and } 0<j<c+N \\ q_{m,j-1} & j = c+N \end{cases} \tag{4}$$

When $m+j < c+N$, the payment area will have $m+j$ payers and will not be full even if no one leaves the payment area until the tagged customer becomes the first in the outside line. This follows since a customer from the m customers in the queue can enter the store only if a shopping customer moves to the payment area. In other words, there is no possibility that the payment area will be full when the tagged customer becomes the first customer in the outside queue. When $j = 0$, the next event is a shopper leaving the shopping area and joining the payment area while a customer in the queue outside enters the store. Thus, $q_{m,j} = q_{m-1,j+1}$. When $j = c+N$, the next event is a customer leaving the payment area since customers cannot move from the shopping area to the payment area when the payment area is full, and $q_{m,j} = q_{m,j-1}$. When $0 < j < c+N$, the next event is either (i) a shopper leaving the shopping area and joining the payment area while a customer from outside enters the store (with probability $a(j)$) or (ii) a payer leaving the payment area (with probability $(1-a(j))$). Note the special case when $m+j = c+N$. In that case, $q_{c+N-j,j}$ is the probability that $c+N-j$ outside customers will enter the shopping area sequentially before a payer leaves the payment area. Thus,

$$q_{m,j} = q_{c+N-j,j} = a(j)\cdot a(j+1)\cdot a(j+2)\cdots a(c+N-1) = \prod_{t=1}^{m} a(c+N-t).$$

3.1. Waiting times

Assume that there are n customers waiting in the outside line and that a tagged customer who finds that $L_1^{OUT} = n, 0 \leqslant n < T$, (i.e., $L_1^{SHOP} = K$ and $L_1 = K+n$) joins the outside line at time t_0 so that L_1^{OUT} becomes $n+1$. Fig. 3 depicts the time instances along the outside queue. For example, when $n = 4$, four customers are already in line when the tagged customer joins the outside line at time t_0 where t_m represents the time at which the m^{th} customer in the outside line enters the store. In other words, t_1 is the time at which the first customer currently in line enters the store, t_2 is the time at which the second customer in line enters the store, and so on. Finally, t_5 is the time at which the tagged customer (who was fifth in line upon arrival) enters the store. Let $W_{m,n} = t_m - t_{m-1}$ $m = 1, \dots, n+1$ be the waiting time of the m^{th} customer in the outside line until that customer becomes the first one in the line given that $L_1^{OUT} = n+1$. Let $W_n^T = \sum_{m=1}^{n+1} W_{m,n}$ $(= t_{n+1} - t_0)$ be the total waiting time of a tagged customer who observes $L_1^{OUT} = n$ and joins the outside queue waiting to enter the store.

Proposition 1.. *Assume that the outside queue has only T waiting places. Let W_n^T be the outside waiting time of a customer who finds $L_1^{OUT} = n \geqslant 0$ (i.e., the system state is $(L_1 = K+n, L_2 = j)$ and joins the line so that L_1^{OUT} becomes $n+1$. Then,*

$$W_n^T = \sum_{m=1}^{n+1} \sum_{j=0}^{c+N} \frac{P_{K+m,j}}{\overline{P}_{K+m}\cdot \overline{e}} \left(q_{m-1,j}(Exp(c\mu) + Exp(K\xi)) + \left(1 - q_{m-1,j}\right) Exp(K\xi) \right)$$

where $\overline{e}$ is a column vector with all its elements equal to 1.

Proof:. *By the preceding definitions, $W_n^T = \sum_{m=1}^{n+1} W_{m,n}$ (see the illustration in Fig. 3). Define $W_{m,n,j} \equiv W_{m,n} | L_2 = j$. If $L_2 = c+N$ when the tagged customer becomes the first person in the outside line (which occurs with probability $q_{m-1,j}$), the tagged customer must wait for a payer to leave the payment area and then for a shopper to move from the shopping area to the payment area (i.e., waiting time $Exp(c\mu) + Exp(K\xi)$). Otherwise, under the complementary probability $1 - q_{m-1,j}$, the customer waits only for a shopper to leave the shopping area (i.e., waiting time $Exp(K\xi)$).*

Thus, $W_{m,n,j} = q_{m-1,j}(Exp(c\mu) + Exp(K\xi)) + (1 - q_{m-1,j})Exp(K\xi)$ and

$$W_{m,n} = \sum_{j=0}^{c+N} P(L_2 = j | L_1 = K+n) \cdot W_{m,n} | L_2 = j$$

$$= \sum_{j=0}^{c+N} P(L_2 = j | L_1 = K+n) \cdot \left(q_{m-1,j}(Exp(c\mu) + Exp(K\xi)) + \left(1 - q_{m-1,j}\right) Exp(K\xi) \right)$$

$$= \sum_{j=0}^{c+N} \frac{P_{K+n,J}}{\overline{P}_{K+n}\cdot \overline{e}} \cdot \left(q_{m-1,j}(Exp(c\mu) + Exp(K\xi)) + \left(1 - q_{m-1,j}\right) Exp(K\xi) \right)$$

Then,

$$W_n^T = \sum_{m=1}^{n+1} W_{m,n} = \sum_{m=1}^{n+1} \sum_{j=0}^{c+N} \frac{P_{K+m,j}}{\overline{P}_{K+m}\cdot \overline{e}} \left(q_{m-1,j}(Exp(c\mu) + Exp(K\xi)) + \left(1 - q_{m-1,j}\right) Exp(K\xi) \right)$$

. ■

By Proposition 1, the expected waiting time of a customer who finds $L_1^{OUT} = n \geqslant 0$ and joins the outside queue is

$$E\left[W_n^T\right] = \sum_{m=1}^{n+1} \sum_{j=0}^{c+N} \frac{P_{K+m,j}}{\overline{P}_{K+m}\cdot \overline{e}} \left(q_{m-1,j} \frac{1}{c\mu} + \frac{1}{K\xi} \right) \tag{5}$$

Corollary 1.. *Limit properties of $E\left[W_n^T\right]$:*

(i) When $\lambda \to \infty$, the shopping area is almost always full (equals K) and the payment area becomes an $M(K\xi)/M(\mu)/c/N$ queue (i.e., c parallel servers and additional N waiting places) with steady state probabilities of $P_j = \frac{a^c}{c!} P_0\ j = 0, \dots, c-1$; $P_{c+k} = \frac{a^c}{c!} \rho^k P_0\ \ k = 0, \dots, N$;

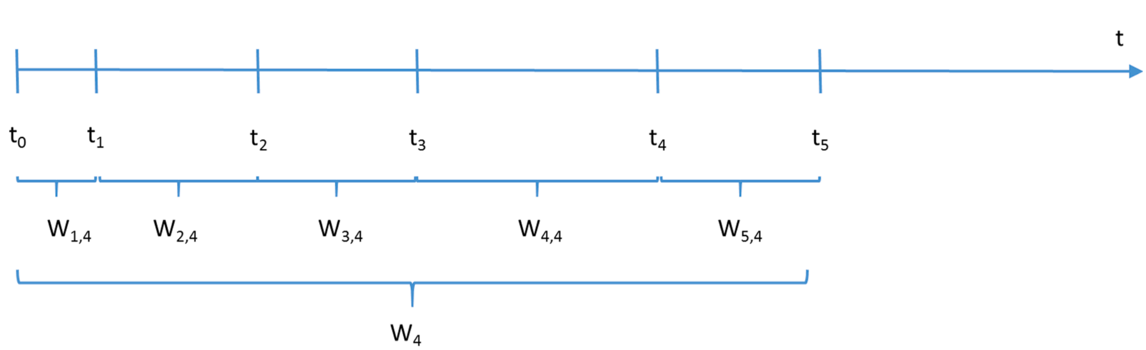

Fig. 3. Time instances along the outside line when a customer who finds $L_1^{OUT} = 4$ joins the queue.

$P_0 = \left[\sum_{n=0}^{c-1}\frac{a^n}{n!} + \frac{a^c}{c!}\sum_{n=c}^{N+c}\rho^{n-c}\right]^{-1}$ where $a = \frac{K\xi}{\mu}$ $\rho = \frac{K\xi}{c\mu}$. Thus, the probability that the payment area is full is $P_{c+N} = \frac{a^c}{c!}\rho^N P_0$. $\lim_{\lambda\to\infty} E\left[W_n^T\right]$

$$= (n+1)\left(P_{c+N}\left(\frac{1}{c\mu}+\frac{1}{K\xi}\right) + (1-P_{c+N})\frac{1}{K\xi}\right)$$
$$= (n+1)\left(P_{c+N}\frac{1}{c\mu}+\frac{1}{K\xi}\right)$$

(ii) When $\lambda\to 0$, the customer who joins the outside queue is the first in line. Thus,

$$\lim_{\lambda\to 0} E\left[W_n^T\right] = \sum_{j=0}^{c-1} P_j\left(\frac{1}{j\mu}+\frac{1}{K\xi}\right) + \sum_{j=c}^{c+N} P_j\left(\frac{1}{c\mu}+\frac{1}{K\xi}\right) = \sum_{j=0}^{c-1} P_j\frac{1}{j\mu} + \sum_{j=c}^{c+N} P_j\frac{1}{c\mu} + \frac{1}{K\xi}$$

(iii) As $\mu\to\infty$, customers do not wait in the payment area. Thus, since the payment area is empty, the shopping area is depleted at a rate of $K\xi$, $\lim_{\mu\to\infty} E\left[W_n^T\right] = (n+1)\frac{1}{K\xi}$.

(iv) As $\xi\to\infty$, customers do not "orbit" in the shopping area. The entire system becomes an $M(\lambda)/M(\mu)/c/N+T$ queue. However, since the payment area is full, $\lim_{\xi\to\infty} E\left[W_n^T\right] = (n+1)\frac{1}{c\mu}$.

(v) Clearly, $\lim_{\mu\to 0} E\left[W_n^T\right] = \lim_{\xi\to 0} E\left[W_n^T\right] = \infty$.

3.2. Outside risk of infection

We define a measure of the risk of infection for a tagged customer waiting outside. The risk is proportional to the total number of other customers the tagged customer meets while queueing outside before entering the store. Let $R_{m,n,j}^T$ be the number of customers that the m^{th} customer meets while waiting in the outside line given that there are n customers outside and j customers at the payment area. That is, the system state is $(L_1 = K+n, L_2 = j)$. Denote by $d^{SHOP}(n,j)$ the probability of moving from state $(L_1 = K+n, L_2 = j)$ to state $(L_1 = K+n-1, L_2 = j+1)$. We use the superscript *SHOP* to indicate that a shopper completes shopping and moves to the payment area, implying that the first customer ($m = 1$) in the outside queue enters the shopping area. Denote by $d^{JOIN}(n,j)$ the probability of moving from state $(L_1 = K+n, L_2 = j)$ to state $(L_1 = K+n+1, L_2 = j)$. In that case, the superscript *JOIN* indicates that an arriving customer joins the outside queue. Denote by $d^{PAY}(n,j)$ the probability of moving from state $(L_1 = K+n, L_2 = j)$ to state $(L_1 = K+n, L_2 = j-1)$, which is the probability that a payer completes paying and leaves the store. These probabilities are extracted from Fig. 2.

Then, the risk of infection for the tagged customer is proportional to

$$R_{m,n,j}^T = \begin{cases} d^{SHOP}(n,j)(n-1) + d^{JOIN}(n,j)R_{m,n+1,j}^T + d^{PAY}(n,j)R_{m,n,j-1}^T & m = 1 \\ d^{SHOP}(n,j)\left(1+R_{m-1,n-1,j+1}^T\right) + d^{JOIN}(n,j)R_{m,n+1,j}^T + d^{PAY}(n,j)R_{m,n,j-1}^T & m > 1 \end{cases} \tag{6}$$

If $m = 1$, the tagged customer is first in the outside line. Thus, with probability $d^{SHOP}(n,j)$, the next event is the tagged customer entering the store and the number of customers the tagged customer met equals $n-1$. With probability $d^{JOIN}(n,j)$, a new customer joins the outside queue so the tagged customer will meet $R_{m,n+1,j}^T$ customers while waiting outside. Finally, with probability $d^{PAY}(n,j)$, a payer completes payment and leaves the store, resulting in $j-1$ customers in the payment area, and the tagged customer will meet $R_{m,n,j-1}^T$. Thus, $R_{m,n,j}^T = d^{SHOP}(n,j)(n-1) + d^{JOIN}(n,j)R_{m,n+1,j}^T + d^{PAY}(n,j)R_{m,n,j-1}^T$.

If $m > 1$, $R_{m,n,j}^T$

$$= d^{SHOP}(n,j)\left(1+R_{m-1,n-1,j+1}^T\right) + d^{JOIN}(n,j)R_{m,n+1,j}^T + d^{PAY}(n,j)R_{m,n,j-1}^T$$

since, with probability $d^{SHOP}(n,j)$, the next event is a customer ahead of the tagged customer entering the store. The number of customers the tagged customer will have met equals 1 (having already met the customer ahead who enters the store), and the tagged customer will then meet $R_{m-1,n-1,j+1}^T$ additional customers in the store. The effect of the other two events (with probabilities $d^{JOIN}(n,j)$, and $d^{PAY}(n,j)$, respectively) is the same as when $m = 1$. Thus, we state,

Proposition 2. *Assume that the outside queue has only T waiting places. Let R_n^T be the number of customers a tagged customer who finds $L_1^{OUT} = n$ (i.e., the system state is $(L_1 = K+n,\ L_2 = j)$ and joins the queue meets while waiting outside so that L_1^{OUT} becomes $n+1$. Then, R_n^T*

$$= \sum_{j=0}^{c+N}\frac{P_{K+n,j}}{\overline{P}_{K+n}\cdot\overline{e}}\ R_{n+1,n+1,j}^T$$

.

By Proposition 2, the tagged customer's expected risk of infection when finding $L_1^{OUT} = n\geqslant 0$ and joining the outside queue is

$$E\left[R_n^T\right] = \sum_{j=0}^{c+N}\frac{P_{K+n,j}}{\overline{P}_{K+n}\cdot\overline{e}}\ E\left[R_{n+1,n+1,j}^T\right] \tag{7}$$

Corollary 2. *Limit properties of $E\left[R_n^T\right]$:*

(i) As $\lambda\to\infty$, the outside queue is almost always equal to the threshold T. Thus, $\lim_{\lambda\to\infty} E\left[R_n^T\right] = n+T-1$. As $\lambda\to 0$, almost no customers arrive after the tagged customer and $\lim_{\lambda\to 0} E[R_n^T] = n$.

(ii) Clearly, $\lim_{\mu\to 0} E\left[R_n^T\right] = \lim_{\xi\to 0} E\left[R_n^T\right] = n+T-1$.

4. Individual optimization

Define U_T^T as the utility of a customer who joins a full queue (size T). By Eq. (1), $U_T^T \equiv r - \gamma W_T^T - \delta R_T^T$. We next find the maximum queue size set by an individual ('selfish') customer. That is, we identify an equilibrium individual threshold strategy T_e that satisfies $E[U_{T_e}^{T_e}] < 0$ and $E[U_{T_e-1}^{T_e}] \geqslant 0$.

For example, consider a store in which the maximum number of allowed shoppers is eleven and the payment area has three cashiers and a maximum adjacent waiting space of two customers (i.e., $K = 11$, $c = 3$, $N = 2$). Set $\lambda = 18$, $\mu = 10$, and $\xi = 3$. First consider a customer whose only concern is the waiting cost so that $\delta = 0$ and utility, U_T^T, is reduced to $U_T^T = r - \gamma W_T^T$. Thus, from Fig. 4a, the optimal individual strategy T_e can be derived. If $\gamma = 1$ and $r = 1$, then, as depicted in Fig. 4a, $E[W_T^T] = 1$ at $T = 24.5$ so the optimal individual threshold is $T_e = 25$. When the

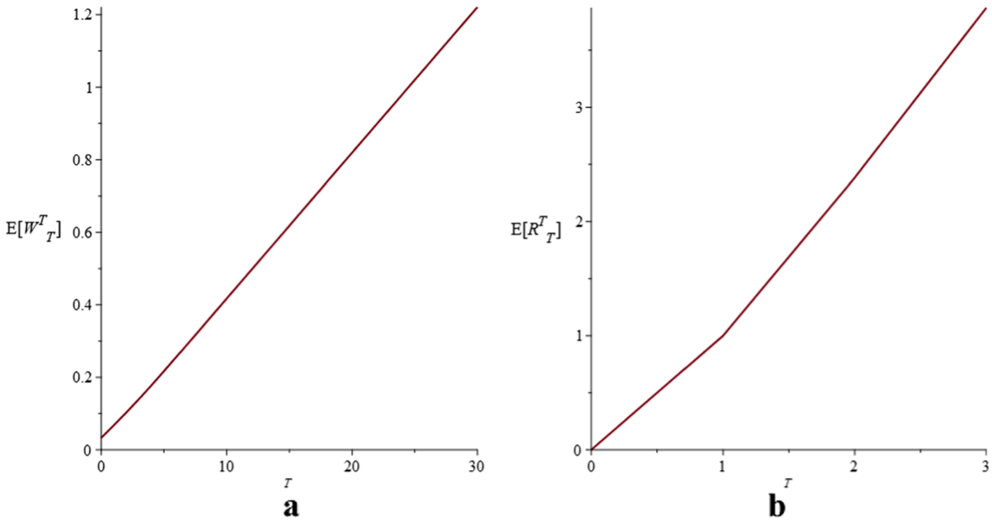

Fig. 4. $E[W_T^T]$ and $E[R_T^T]$ as a function of T when $\lambda = 18,\ \mu = 10,\ \text{and}\ \xi = 3$.

customer's only concern is risk of infection (i.e., $\gamma = 0$), $U_T^T = r - \delta R_T^T$. For that case, when $\delta = 1$ and $r = 1$, Fig. 4b implies that $T_e = 2$ as it is considerably lower than the threshold when only waiting time is taken into account.

Next, we perform a sensitivity analysis to investigate the effect of changing the parameter values on the optimal individual strategy T_e. We start with the case of a customer whose only concern is the waiting cost so utility is $U_T^T = r - \gamma W_T^T$. For example, when $r = 1$ and $\gamma = 10$, Fig. 5 depicts the optimal strategy T_e as a function of model parameters λ, μ, and ξ, respectively.

Now, consider the case of a customer whose utility is $U_n^T = r - \delta R_n^T$. That is, a customer who is concerned only about risk of infection. When, for example, $r = 1$ and $\delta = 0.1$, Fig. 6 depicts the optimal strategy T_e as a function of the model parameters.

In contrast to Naor's (1969) model, our optimal individual strategy depends on the arrival rate λ. However, when the arrival rate is relatively high ($\lambda > 15$ in Fig. 5, $\lambda > 16$ in Fig. 6), the threshold remains constant for any value of λ. This follows since, under such an arrival rate, the store becomes highly congested and the threshold is not sensitive to a change in λ. Moreover, a customer whose concern is mainly about risk of infection will have a strategy that is more sensitive to changes in λ than a customer who considers only the cost of waiting outside. This result follows since risk of infection considers also the number of customers who arrive after the tagged customer while waiting time depends only on the number of customers ahead. As depicted in Figs. 5 and 6, increasing ξ (shopping rate) or μ (service rate) increases the optimal threshold of an individual exercising the T-strategy. It follows that the strategy of a customer whose concern is reducing waiting time will be more sensitive to changes in ξ or μ because increasing ξ or μ allows customers ahead of the tagged customer to enter the store sooner, thus reducing the tagged customer's waiting time. However, a change in ξ or μ only partially helps a customer who considers only infection risk since the risk further depends on customers who will arrive after the tagged customer.

5. Overall optimization

Now we identify the maximum queue size T_{So} set by the authority that maximizes overall social welfare, which is defined as the total expected net benefit for all customers per unit of time. Denote by U^T the utility of an arriving customer who joins a full queue of size T (given a maximum queue size of T). The probability of observing T customers in the outside queue is $\overline{P}_{K+T}\cdot\overline{e}$. Thus, the probability that a customer will join equals $\left(1 - \overline{P}_{K+T}\cdot\overline{e}\right)$ and

$$E[U^T] = \frac{\sum_{i=0}^{K-1}\overline{P}_i\cdot\overline{e}\cdot r + \sum_{n=0}^{T-1}\overline{P}_{K+n}\cdot\overline{e}\cdot E[U_n^T]}{\left(1 - \overline{P}_{K+T}\cdot\overline{e}\right)} \quad (8)$$

The first summation term in the numerator of Eq. (8) relates to the probability that an arriving customer will enter the store immediately upon arrival so there is no waiting cost or infection-risk cost involved.

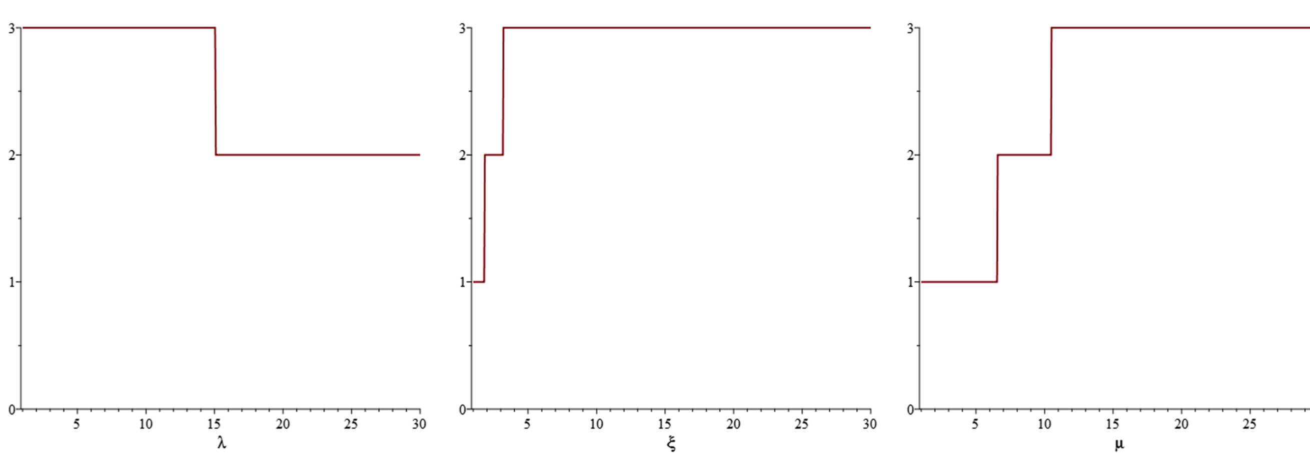

Fig. 5. T_e as a function of model parameters when $r = 1,\ \delta = 0,$ and $\gamma = 10$ (considering only waiting).

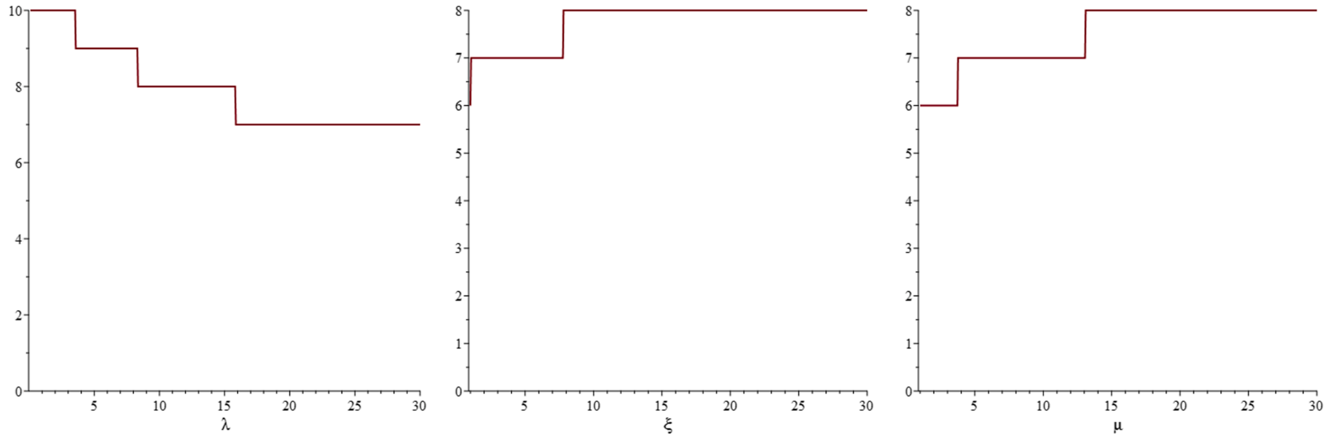

Fig. 6. T_e as a function of model parameters when $r = 1$, $\delta = 0.1$, and $\gamma = 0$. (considering only risk).

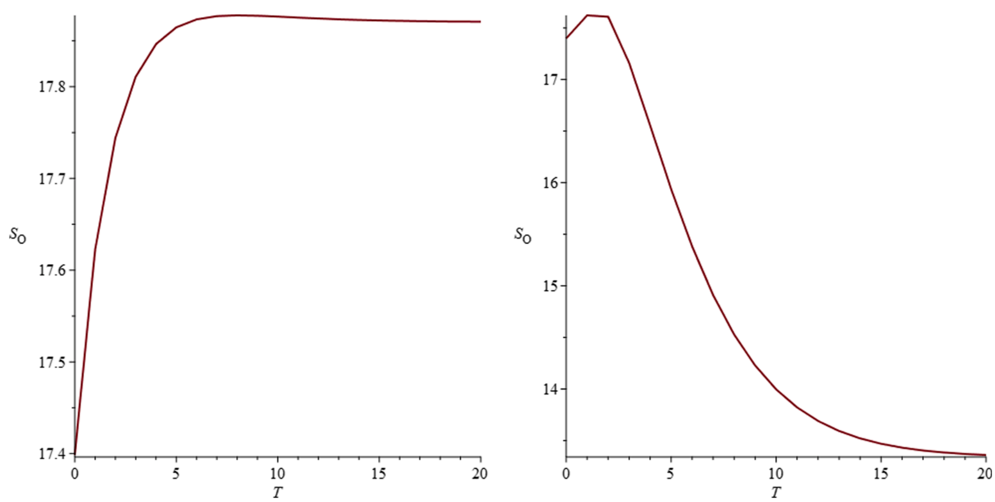

Fig. 7. S_O as a function of T. **7a** r = 1, γ = 1, δ = 0 (considering only waiting) , **7b** r = 1, γ = 0, δ = 1 (considering only risk).

The second term deals with customers who wait outside and incur both a waiting cost and an infection-risk cost and where $E[U_n^T] \equiv r - \gamma E[W_n^T] - \delta E[R_n^T]$.

Denote by S_O the expected social benefit per unit of time. Thus,

$$S_O = \lambda(1 - \overline{P}_{K+T}\cdot\overline{e})E[U^T] \equiv \lambda_{eff}E[U^T] \quad (9)$$

That is, T_{So} maximizes S_O in Eq. (9).

Using the same numerical example as before – $\lambda = 18$, $\mu = 10$, $\xi = 3$, $K = 11$, $c = 3$, $N = 2$ – and setting $r = 1$, Fig. 7 depicts S_O as a function of T for different values of γ and δ.

As Fig. 7a shows, when a customer's only concern is the waiting cost, the optimal social threshold is $T_{So} = 8$. However, when a customer's

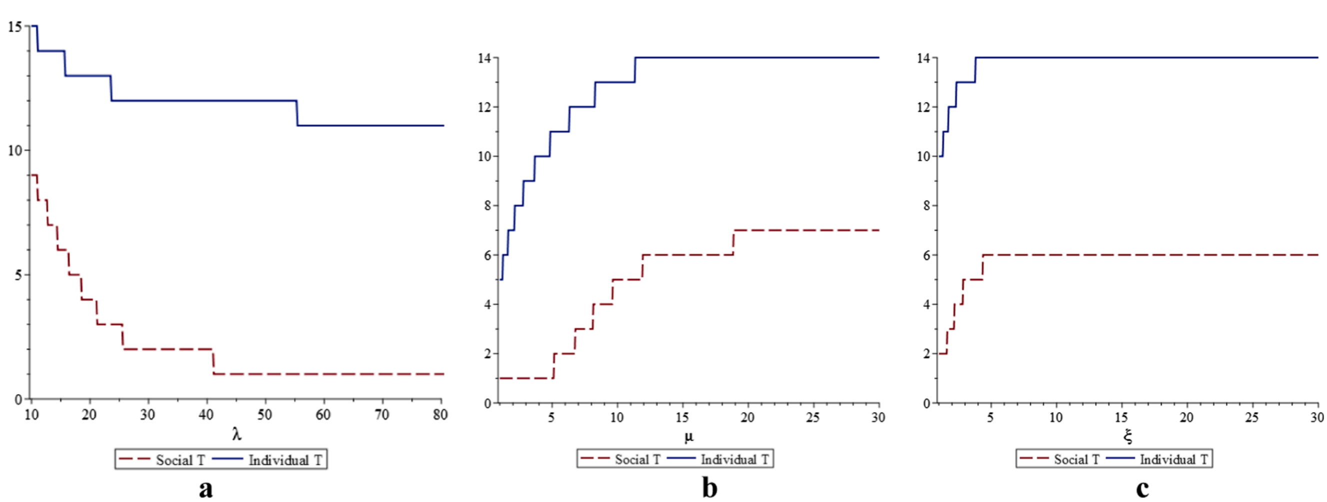

Fig. 8. T_{So} and T_e as a function of model parameters when $r = 1$, $\gamma = 0.4$, and $\gamma = 0.04$.

only concern is infection risk, Fig. 7b, the optimal social threshold is $T_{So} = 1$.

The effects of changing the parameter values on the optimal social and individual strategies are depicted in Fig. 8.

Fig. 8 shows that both optimal thresholds decrease in λ and increase in μ and ξ. Furthermore, the figure illustrates the gap between the optimal thresholds under individual versus social strategies. As in Naor's (1969) model, the optimal social threshold (the maximal queue size set by the authority) is always smaller than the optimal individual threshold (the maximal queue size set by an individual customer). Hanukov and Yechiali (2020) present a case in which this relationship does not necessarily hold.

6. Conclusions

The risk of infection by the COVID-19 virus has influenced the strategic behavior of customers who have to wait in line outside stores to obtain service. Consequently, a customer's decision whether to join such a queue or balk, depends both on the expected waiting time in line as well as on the risk of infection, which is proportional to the total number of other customers who will share the queue, including those already in line and those joining prior to the customer entering the premises. We derive expressions for mean waiting times and infection risk and obtain explicit formulas for limit values of the parameters. We construct a novel reward-cost-risk utility structure that determines customers' strategic behavior. Customer decisions in our model depend not only on the observed length of the outside queue but also on the unobserved number of customers in the payment area. Moreover, it is shown that an individual customer's strategy depends in part on the actions of other customers. Numerical examples calculate the threshold values as a function of the system's parameters.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Alban, A., Chick, S.E., Dongelmans, D.A., van der Sluijs, A.F., Wiersinga, W.J., Vlaar, A. P., Sent, D., 2020. ICU capacity management during the COVID-19 pandemic using a stochastic process simulation.

Avrachenkov, K., Nain, P., Yechiali, U., 2014. A retrial system with two input streams and two orbit queues. Queueing Systems 77 (1), 1–31.

Hanukov, G., Yechiali, U., 2020. Customers strategic behavior in a two-stage service system with servers' vacations where customers can use a time-saving APP. Working paper.

Hassin, R., 2016. Rational queueing. CRC Press.

Hassin, R., Haviv, M., 2003. To queue or not to queue: Equilibrium behavior in queueing systems, Vol. 59. Springer Science & Business Media.

Kaplan, E.H., 2020. OM forum—COVID-19 scratch models to support local decisions. Manuf. & Service Oper. Manage. 22 (4), 645–655.

Long, D.Z., Wang, R., Zhang, Z., 2020. Pooling and balking: Decisions on COVID-19 testing. Available at SSRN 3628484.

Naor, P., 1969. The regulation of queue size by levying tolls. Econometrica: J. Econ. Soc. 15–24.

Perel, N., Yechiali, U., 2014. The Israeli queue with retrials. Queueing Syst. 78 (1), 31–56.

Perlman, Y., Yechiali, U., 2020a. On tandem stochastic networks with time-deteriorating product quality. Int. J. Prod. Res. 58 (13), 3956–3964.

Perlman, Y., Yechiali, U., 2020b. Reducing risk of infection – the COVID-19 queueing game. Saf. Sci. 132, 104987.

Yechiali, U., 1971. On optimal balking rules and toll charges in the GI/M/1 queue. Oper. Res. 19, 349–370.

Yechiali, U., 1972. Customers' optimal joining rules for the GI/M/s queue. Manage. Sci. 18, 434–443.