Omega 125 (2024) 103024

Contents lists available at ScienceDirect

Omega

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

Mail back or in-store dropoff? Optimal design of product-exchange policies in omnichannel retailing systems[☆]

Ke Sun [a], Yunan Liu [b,*], Xiang Li [c]

[a] School of Economics and Management, Beijing University of Chemical Technology, Beijing, 100029, China
[b] Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, NC 27695-7906, United States of America
[c] School of Economics and Management, Chang'an University, Xi'an, 710064, China

ARTICLE INFO

Keywords:
Queueing theory
Queueing economics model
Retailing system
Omnichannel retailing
Product exchange policy

ABSTRACT

Online retailing has been booming over the past several years as people grow increasingly comfortable with it and accustomed to its ease and speed. A major drawback of online retailing is the lack of consumer–product interaction before a purchase is finalized, which often leads to consumer dissatisfaction due to mismatched expectation of the received product. In response, retailers usually promise that online orders may be returned or exchanged free of extra charges, which can be processed either online (e.g., by mail) or onsite i.e., in-store dropoff. In this work, we study the implications of both online and onsite exchange policies in the setting of a queueing model that offers omnichannel services. Taking into account customers' behavioral responses to these policies, we aim to inform the retailer of the one that generates a higher system revenue. Our results reveal that the online exchange policy is a double-edged sword: On the one hand, it helps eliminate the inconvenience cost for exchange customers to revisit the store; on the other hand, it can trigger more feedback orders and render a higher system congestion level, which in turn, deters future customers from placing orders. Specifically, we discover that online exchange becomes an inferior policy relative to in-store exchange when the market size is large.

1. Introduction

Online retailing has been booming in recent years due to the prevalence of smartphones and access to the Internet. Top online shopping categories include fashion, entertainment, electronics, food, etc. The main advantage of online shopping is its convenience, because consumers no longer need to travel to the physical store. Besides, it brings additional benefits during the Covid-19 pandemic by largely reducing the risk of infection (it avoids close contacts to other consumers in a physical store). Indeed, many retailers are offering both online and in-person shopping (i.e., multichannel or omnichannel retailing), see [1–4].

However, some things are unique to an in-person experience such as physically checking and testing a product to ensure its quality. For perfumes and cosmetics, stores offer the opportunity to test the products to make sure they are right for the consumers; it is also nice to try on clothing at the store in order to personally choose the best style and fit. Hence, a major drawback of online shopping, despite its many benefits, is the lack of consumer–product interaction before a purchase is finalized, which may lead to consumer dissatisfaction due to mismatched expectation of the received products. To ensure the high competitiveness of online shopping, omnichannel retailers usually promise that online orders, if not meeting customers' expectations, may be returned or exchanged free of extra charges [5–7].

In practice, two policies have been widely used for product exchange and return: (i) by mail (online) and (ii) in-store dropoff (onsite). For example, *Best Buy* allows orders purchased online to be returned or exchanged either by mail or in-store [8]; *Old Navy* only accepts returns and exchanges in store [9]; while *eBay* requires that customers ship the product to the seller by mail [10]. In this paper, we are motivated to study an omnichannel retailing system that offers both online and in-person shopping. Because online shopping induces feedback orders (consumers demanding return/exchange when receiving products that do not meet their expectations), we investigate the impact of the two aforementioned policies on the treatment of these feedback orders. Hereby we only study product exchange and do not consider product return. We hope to answer the following questions: (1) How do different exchange policies shape customers' behavior and experience? (2) From the service provider's perspective, which exchange policy can generate a higher revenue?

☆ Area: Production Management, Scheduling and Logistics. This manuscript was processed by Associate Editor Harks.

* Corresponding author.
E-mail addresses: kesun@buct.edu.cn (K. Sun), yliu48@ncsu.edu (Y. Liu), lixiang@chd.edu.cn (X. Li).

https://doi.org/10.1016/j.omega.2023.103024
Received 8 December 2022; Accepted 11 November 2023
Available online 17 January 2024
0305-0483/

We develop a new queueing economics model in which the service facility is able to process orders received in two channels: online and onsite. All arriving orders are buffered in a common queue and processed by a single server according to the *first-come first-served* (FCFS) service discipline. An online order, upon completion, may or may not meet the expectation of its corresponding customer, and if not, will be fed back to the service facility for a remake (i.e., a product exchange occurs). Exchange orders are treated equally as new orders (they join the same queue and are processed under FCFS). Moreover, customers are delay sensitive and strategic; they make two-step decisions. Each arriving customer will first decide whether to place an order (i.e., join the service system), and if yes, she will next select the "right" channel to place an order (online or onsite). The onsite channel guarantees the service (or product) to be well suited to a customer but incurs an inconvenience cost; while the online channel spares the above-mentioned inconvenience cost but may generate excessive delay in the future in case customers revisit the service system for exchange.

Our contributions are summarized below.

- **Equilibrium strategy.** We study two omnichannel retailing queue models that are distinct solely in the product exchange policies: one allows only online exchange (dubbed the *online exchange model*) and the other allows only onsite exchange (dubbed the *onsite exchange model*). In both models, we derive customers' joint order-placing and channel-selecting equilibrium strategies.
- **Throughput and revenue.** Using the equilibrium strategy, we compute the system throughput in both models with an exogenous service fee (i.e., which is a fixed parameter). We later allow the service provider to adjust the service fee in order to maximize his revenue; we compare the optimal revenue under the two exchange policies.
- **Findings and insights.** Our results reveal several interesting and insightful findings. Contrary to the consensus that online exchange is in general more beneficial, we discover that the online-exchange model loses its advantage when the system is highly congested and the service fee is low. In addition, under the setting of endogenous pricing (where the service provider can adjust the service fee in order to maximize his revenue), online exchange is a less profitable policy when the system is highly congested and the risk of exchange is intermediate. Another interesting result is that the system throughput is not always increasing when the risk of exchange decreases. Our conclusions are drawn from theoretical results, numerical experiments, and in-depth discussions.

The remainder of the article is organized as follows. In Section 2, we review the most related literature. Section 3 sets up an omnichannel retailing queueing model. In Section 4, we characterize the joint equilibria in both online-exchange and onsite-exchange models. In Section 5, we investigate two extensions of our base model. First, in Section 5.1 we study the setting of endogenous service price. Next, in Section 5.2, we consider the case in which customers' inconvenience costs is considered to follow (uniform) random distribution. Finally, in Section 6, we draw conclusions and discuss future research directions. All proofs are provided in Appendix.

2. Related literature

Three lines of research are most relevant to the present study.

Our paper contributes to the economic analysis of queueing systems, which is pioneered by Naor [11], who studies an observable queueing system where customers rationally decide whether to join based on observed queue information. Edelson and Hilderbrand [12] extend this problem to the unobservable setting. These two seminal works have inspired numerous works on the queueing economics with strategic consumers. For comprehensive reviews of the related works, see [13, 14]. Also see recent survey by Economou [15] on the effect of information structure on strategic customer behavior and the psychological principle on them [16] In particular, the present paper is relevant to queueing-economics models with a two-dimensional customer strategy. Hassin and Roet-Green [17] consider an unobservable queueing model where arrivals are able to pay for inspecting the queue length information; and based on this information, they make their joining decision. Hassin and Roet-Green [18] investigate an order-onsite model in which customers first decide whether to travel to the store and then whether to join the queue upon arrival. Sun et al. [19] study an order-ahead model in which customers first order online and then travel to the service facility. In addition, the profit optimization via pricing constitutes a prevalent research topic within online service models, see [20,21], etc.

Distinct from the above literature, the present work studies a queueing-economics model having customer feedback (hereby the exchange customers). In addition, different from [18,19], in our model customers are given the option to select the channel (either online or onsite) for receiving their services.

This work is also related to recent developments on omnichannel service systems. [22] investigate an omnichannel retailing model that allows consumers to pay online and pick up in store. Later, they extend to the setting of a restaurant model [23] and show that the online order-placing technology can help reduce the delay for both tech-savvy (who place orders remotely) and conventional (who place orders in store) customers. Multiclass queueing systems have been developed to determine the optimal staffing [24] and scheduling [25] rules for the omnichannel customer contact centers. [26] study how the level of information in omnichannel service models impacts the system performance. [27] consider a restaurant model with a food delivery platform that can serve two streams of customers: tech-savvy customers and conventional customers. Their findings reveal that the restaurant's revenue may decline as more customers join the online service channel. [28] suggest that the advent of online-to-offline (O2O) platforms affords traditional retail outlets the opportunity to broaden their customer base via the integration of online ordering and delivery services. We also refer readers to a comprehensive review [29]. Our work is partially inspired by Baron et al. [30] which study a queueing model with an additional online order-placing option; they show that this new option improves the system throughput but is detrimental to both consumers' individual utilities and the social welfare. In contrast to the papers above, our work focuses on the design of the product exchange policy in omnichannel service models rather than studying the benefit of omnichannel services. In addition, the focus of the present work is to investigate the omnichannel aspect in the post-service stage, which is another major distinction from the above literature.

Last, our paper is related to the retailing literature on models with customer return and exchange. Retailing systems with exchange orders are often modeled as queues with customer feedback. Guo et al. [31] investigate the strategic customers' behavior in queueing model with feedback. Aziz and Wahid [32] find that non-online shoppers prefer to shop traditionally than online shopping because they would like to self-assess the quality of the products. [33] show that effective return and exchange policies help incentivize customers to revisit the system for more services. [34] study an online retailing model in which the customers receiving unsatisfied products make endogenous exchange/return decisions. They find that, comparing to product exchange, product return yields a lower social welfare. [35] study the pricing and return policy decisions in an omnichannel firm, where customers are offered two options: buy online and return online and buy online and return in store. The present paper draws distinctions from the above literature by studying the design of product-exchange policies in face of omnichannel services.

3. Model description

We consider a single-server queueing model having potential arrivals in accordance with a Poisson process with rate Λ (referred to as *potential market size*), and *independent and identically distributed* (I.I.D.) service times following an exponential distribution with rate μ (referred to as *service capacity*). The arrival times and service times are independent. Customers are delay sensitive and incur a cost C per unit of time in the system. The system offers omnichannel services: each arriving customer may choose to either place an online order or an onsite order:

- **Onsite orders:** Customers placing orders in the onsite channel incur an inconvenience cost $C_h > 0$. (Here C_h roughly represents the inconvenience of traveling to the store, inflexibility of shopping window, risk of getting infected during the pandemic, or a mixture of all.) On the other hand, we assume that all services completed onsite will perfectly match with customers' expectation (because customers can have sufficient personal interactions with the products during their visit to the store), so there is no need for exchange. We first treat C_h as the same constant for all customers and later extend to the case where C_h is a random variable in order to capture the heterogeneity in customers' inconvenience cost; see Section 5.2.
- **Online orders:** Orders placed in the online channel will not incur the inconvenience cost C_h. However, received products may or may not meet the customers' expectation, which in turn may trigger after-sales product exchanges (i.e., feedback orders). A customer is satisfied with the online service with probability $\alpha \in (0,1)$. We assume a customer is able to gain sufficient knowledge from the first service experience (even if the product is dissatisfying and an exchange is requested) so that the new product/service is guaranteed to satisfy the customer's need. We hereby refer to α as the *expectation-meeting probability*.

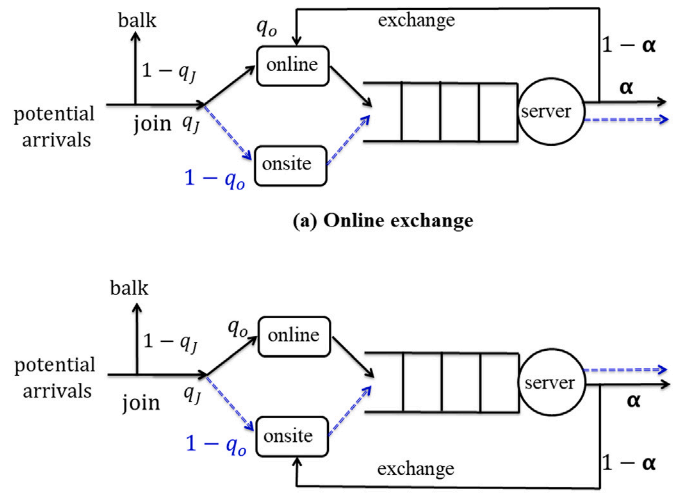
In practice, the online exchange option might also incur a cost $C_o > 0$. For example, an exchange customer may have to leave the package at a drop-off location or post office. Nevertheless, this should be much smaller than the cost of undergoing the entire service process at the onsite store. For simplicity, we normalize this cost to 0 for the online exchange option, and thus, the cost C_h can be regarded as the additional effort cost a customer has to pay when exchanging online rather than onsite.

Orders of all types (online and onsite, exchange and new) are buffered in a common queue and processed under FCFS. All customers pay a service fee P upon placing an order and receive a service reward V upon completing service. (In case exchange is needed, the service reward V is received after the exchange is completed.)

Each potential customer makes decisions in two stages: first, she decides whether *to join or to balk*; second, if joining, whether to *order online or onsite*. We hereby only consider the symmetric equilibrium. We assume that customers do not have the queue length information (i.e., number of pending orders) when they make decisions so their strategy can be described by an *order-placing* probability $q_J \in [0,1]$ and a *channel-selecting* probability $q_o \in [0,1]$, i.e., the probability of ordering online (in-store) is q_o $(1-q_o)$. The assumption of unobservable queue can be justified by the practices of many service systems. For example, many quick-service restaurants (e.g., Starbucks) allow customers to place an order via their mobile apps, which reveal no real-time delay estimation. In addition, their onsite customers are unable to observe precise delay information as well because the order queue consists of "invisible" online orders which may either not have arrived or be waiting somewhere outside the store (e.g., for social distancing during the pandemic).

Given the two-dimensional strategy (q_J, q_o), the *effective external arrival rate* is $\Lambda_e = \Lambda q_J$. By flow conservation, we write out the *effective internal arrival* rate as

$$\lambda_e = \Lambda_e[1+(1-\alpha)q_o] = \Lambda q_J[1+(1-\alpha)q_o].$$



(a) Online exchange

(b) Onsite exchange

Fig. 1. A schematic representation for the omnichannel retailing queue under two product exchange policies: (a) online exchange, and (b) onsite exchange.

Hence, customers' steady-state mean sojourn time (i.e., waiting time plus service time) is

$$W(q_J,q_o) = \frac{1}{\mu-\lambda_e} = \frac{1}{\mu-\Lambda q_J[1+(1-\alpha)q_o]}, \quad \text{for } q_J, q_o \in [0,1], \tag{1}$$

provided that the system is stable, that is, $\lambda_e = \Lambda q_J[1+(1-\alpha)q_o] < \mu$. See Fig. 1 for a schematic illustration of the system's structure and a customer's decision process. All notations are summarized in Table 1.

4. Customers' equilibrium strategies

In this section, we study customers' equilibrium strategies in two models operated under distinct exchange policies: (i) onsite (i.e., in-store dropoff) and (ii) online (i.e., by-mail). We compare the equilibrium performance in the two models.

4.1. Onsite exchange model

In this subsection, we treat the model operated under the onsite-exchange policy. We append a superscript "s" to all notation to indicate that only onsite exchange is allowed. For an online customer, let U_1^s (U_2^s) be the expected utility given that the customer is satisfied (dissatisfied) with the service. In particular, for $q_J, q_o \in [0,1]$, we have

$$U_1^s(q_J,q_o) = V - P - CW(q_J,q_o), \tag{2}$$

$$U_2^s(q_J,q_o) = V - P - C_h - 2CW(q_J,q_o), \tag{3}$$

where the second equality holds because the customer has to undergo the entire queueing process once again. Then the expected utility of an arbitrary online customer (indicated by a subscript "o") is

$$U_o^s(q_J,q_o) = \alpha U_1^s(q_J,q_o) + (1-\alpha)U_2^s(q_J,q_o) = V - P - (1-\alpha)C_h - (2-\alpha)CW(q_J,q_o). \tag{4}$$

On the other hand, the expected utility of an onsite customer (indicated by a subscript "s") is

$$U_s^s(q_J,q_o) = V - P - C_h - CW(q_J,q_o). \tag{5}$$

Remark 1 (*Immediate Comfort vs. Future Convenience*)**.** While online service exhibits an immediate advantage by avoiding the cost C_h, it is possible that customers have to experience an additional delay in case the product/service is unsatisfying. On the other hand, onsite service

Table 1
Glossary of main notation.

Symbol	Definition
μ	Service capacity/rate
Λ	Market size
V	Service reward
C	Delay cost per unit of time
P	Service fee
C_h	Inconvenience cost of onsite service
α	Expectation-matching probability of online service
q_J, q_o	Order-placing probability & online-channel-selecting probability
$W(q_J, q_o)$	Customers' average sojourn time given q_J and q_o
$U_o^s(q_J, q_o), U_s^s(q_J, q_o)$	Expected utility of online and onsite channel under onsite exchange policy
$U_o^o(q_J, q_o), U_s^o(q_J, q_o)$	Expected utility of online and onsite channel under online exchange policy

ensures that the service meets a customer's expectation (thus precluding additional future delay) at the expense of the cost of C_h. Hence, a customer's perspective of online and in-store services is essentially a matter of how to trade-off between *taking a chance to enjoy an immediate comfort* and *staying cautious to reduce the risk of future inconvenience.*

Lemma 1. *Under the onsite exchange policy, for a fixed order-placing probability $q_J \in (0,1]$, there exists a mixed strategy $q_o^s(q_J)$ satisfying*

$$q_o^s(q_J) = \begin{cases} 0, & \text{if } \frac{1}{\eta} < \frac{1}{\mu - \Lambda q_J}, \\ \frac{1}{1-\alpha}\left(\frac{\mu-\eta}{\Lambda q_J} - 1\right), & \text{if } \frac{1}{\mu-\Lambda q_J} \le \frac{1}{\eta} \le \frac{1}{\mu - \Lambda q_J(2-\alpha)}, \\ 1, & \text{if } \frac{1}{\eta} > \frac{1}{\mu - \Lambda q_J (2-\alpha)}, \end{cases} \tag{6}$$

where $\eta = \frac{(1-\alpha)C}{\alpha C_h}$. In addition, $q_o^s(q_J)$ is weakly decreasing in q_J.

Please refer to Appendix for the proof.

Remark 2 (*Onsite Service Is Popular When Congestion Is High*)**.** When q_J increases, the system becomes more congested (with prolonged delay) as more potential customers might choose to join for service. Because there is a chance that online orders may be fed back to endure a second-round queueing process (incurring a cost C_h), some online customers (these who are unsatisfied with the product) are more significantly influenced by the delay increment due to the increase of q_J, which drives more customers to pursue onsite service in the first place.

For a given order-placing probability $q_J \in [0,1]$, let $U_J^s(q_J, q_o)$ be the expected utility of a joining customer. We have

$$U_J^s(q_J, q_o) = q_o^s(q_J) U_o^s(q_J, q_o^s(q_J)) + (1 - q_o^s(q_J)) U_s^s(q_J, q_o^s(q_J)).$$

We next characterize the equilibrium strategy by (q_J^s, q_o^s) in the following theorem. In all cases with $q_J^s = 0$, we say that the equilibrium is $(q_J^s, q_o^s) = \mathbf{0}$ denoting an empty system with no one willing to join for service.

Theorem 1 (*Equilibrium Strategy Under Onsite Exchange*)**.** *Under the onsite exchange policy, we give the equilibrium strategy in three cases specified by the market size Λ.*

a. ***Small market size.*** *If $\Lambda < \mu/(2-\alpha)$, the equilibrium strategy is*

$$(q_J^s, q_o^s) = \begin{cases} \mathbf{0}, & \text{if } \frac{1}{\mu} \ge \max\left\{\frac{V-P-C_h}{C}, \frac{\bar{V}-P}{(2-\alpha)C}\right\}, \\ (\widetilde{q}_J^s, 0), & \text{if } \max\left\{\frac{1}{\mu}, \frac{1}{\eta}\right\} < \frac{V-P-C_h}{C} < \frac{1}{\mu-\Lambda}, \\ (1,0), & \text{if } \frac{1}{\eta} < \frac{1}{\mu-\Lambda} \le \frac{V-P-C_h}{C}, \\ (1, \widehat{q}_o^s), & \text{if } \frac{1}{\mu-\Lambda} \le \frac{1}{\eta} \le \min\left\{\frac{1}{\mu-(2-\alpha)\Lambda}, \frac{(V-P)\alpha}{C}\right\}, \\ (\widehat{q}_J^s, 1), & \text{if } \frac{1}{\mu} < \frac{\bar{V}-P}{(2-\alpha)C} < \min\left\{\frac{1}{\mu-(2-\alpha)\Lambda}, \frac{1}{\eta}\right\}, \\ (1,1), & \text{if } \frac{1}{\mu-(2-\alpha)\Lambda} \le \frac{\bar{V}-P}{(2-\alpha)C} \ \& \ \frac{1}{\mu-(2-\alpha)\Lambda} < \frac{1}{\eta}. \end{cases} \tag{7}$$

b. ***Medium market size.*** *If $\mu/(2-\alpha) \le \Lambda < \mu$, the equilibrium strategy is*

$$(q_J^s, q_o^s) = \begin{cases} \mathbf{0}, & \text{if } \frac{1}{\mu} \ge \max\left\{\frac{V-P-C_h}{C}, \frac{\bar{V}-P}{(2-\alpha)C}\right\}, \\ (\widetilde{q}_J^s, 0), & \text{if } \max\left\{\frac{1}{\mu}, \frac{1}{\eta}\right\} < \frac{V-P-C_h}{C} < \frac{1}{\mu-\Lambda}, \\ (1,0), & \text{if } \frac{1}{\eta} < \frac{1}{\mu-\Lambda} \le \frac{V-P-C_h}{C}, \\ (1, \widehat{q}_o^s), & \text{if } \frac{1}{\mu-\Lambda} \le \frac{1}{\eta} \le \frac{(V-P)\alpha}{C}, \\ (\widehat{q}_J^s, 1), & \text{if } \frac{1}{\mu} < \frac{\bar{V}-P}{(2-\alpha)C} < \frac{1}{\eta}. \end{cases} \tag{8}$$

c. ***Large market size.*** *If $\Lambda \ge \mu$, the equilibrium strategy is*

$$(q_J^s, q_o^s) = \begin{cases} \mathbf{0}, & \text{if } \frac{1}{\mu} \ge \max\left\{\frac{V-P-C_h}{C}, \frac{\bar{V}-P}{(2-\alpha)C}\right\}, \\ (\widetilde{q}_J^s, 0), & \text{if } \max\left\{\frac{1}{\mu}, \frac{1}{\eta}\right\} < \frac{V-P-C_h}{C}, \\ (\widehat{q}_J^s, 1), & \text{if } \frac{1}{\mu} < \frac{\bar{V}-P}{(2-\alpha)C} < \frac{1}{\eta}. \end{cases} \tag{9}$$

where $\widetilde{q}_J^s = \frac{\mu - \frac{C}{V-P-C_h}}{\Lambda}$, $\widehat{q}_o^s = \frac{\mu-\Lambda-\eta}{(1-\alpha)\Lambda}$, $\widehat{q}_J^s = \frac{\mu - \frac{(2-\alpha)C}{\bar{V}-P}}{(2-\alpha)\Lambda}$, $\eta = \frac{(1-\alpha)C}{\alpha C_h}$ and $\bar{V} = V - (1-\alpha)C_h$.

Please refer to Appendix for the proof.

Customers' strategy depends on the arrival rate and service rate of the system. As the arrival rate increases in comparison to the service rate, the queue length increases and thus the waiting time increases. Therefore, fewer customers choose to order. Despite the complex structure of the equilibrium strategy, Theorem 1 exhibits the following trend: As the market size increases, a smaller fraction of arrivals are willing to join for service, among which, even fewer are willing to place online orders. For example, the all-join-all-online case (i.e., $(q_J, q_o) = (1,1)$) is possible only when the market size is small (as in Case (a)). We will conduct numerical experiments in Section 4.3 to gain additional insights of the equilibrium strategy.

4.2. Online exchange model

In this subsection, we focus on the model that is operated under the online-exchange policy. We append a superscript "o" to all notation to indicate that only online exchange is allowed. Consider an online customer, let U_1^o (U_2^o) be her expected utility given that the customer is satisfied (dissatisfied) with the service. For given $q_J, q_o \in [0,1]$, we have

$$U_1^o(q_J, q_o) = V - P - CW(q_J, q_o), \tag{10}$$

$$U_2^o(q_J, q_o) = V - P - 2CW(q_J, q_o), \tag{11}$$

where $W(q_J, q_o)$ is the average sojourn time given in (1), and $U_2^o(q_J, q_o)$ differs from $U_2^s(q_J, q_o)$ in (3) by missing the cost C_h. The expected utility of an online customer is

$$U_o^o(q_J, q_o) = \alpha U_1^o(q_J, q_o) + (1-\alpha) U_2^o(q_J, q_o) = V - P - (2-\alpha) C W(q_J, q_o). \tag{12}$$

The expected utility of an onsite customer (indicated by a subscript “s”) is

$$U_s^o(q_J, q_o) = V - P - C_h - CW(q_J, q_o). \tag{13}$$

Lemma 2. *Under the online exchange policy, for a fixed order-placing probability $q_J \in (0, 1]$, there exists a mixed strategy $q_o^o(q_J)$ satisfying*

$$q_o^o(q_J) = \begin{cases} 0, & \text{if } \frac{1}{\delta} < \frac{1}{\mu - \Lambda q_J}, \\ \frac{1}{1-\alpha}\left(\frac{\mu-\delta}{\Lambda q_J} - 1\right), & \text{if } \frac{1}{\mu-\Lambda q_J} \le \frac{1}{\delta} \le \frac{1}{\mu - \Lambda q_J(2-\alpha)}, \\ 1, & \text{if } \frac{1}{\delta} > \frac{1}{\mu-\Lambda q_J(2-\alpha)}, \end{cases} \tag{14}$$

where $\delta = \frac{(1-\alpha)C}{C_h}$. In addition, $q_o^o(q_J)$ is weakly decreasing in q_J.

Please refer to Appendix for the proof.

The intuition of Lemma 2 is similar to that of Lemma 1. Paralleling Theorem 1, below we characterize the equilibrium structure by (q_J^o, q_o^o). Similar to the onsite exchange model, we hereby organize all results with $q_J^o = 0$ into one case denoted as $(q_J^o, q_o^o) = \mathbf{0}$, meaning that no one is willing to join for service.

Theorem 2 (*Equilibrium Strategy Under Online Exchange*)**.** *Under the online exchange policy, the equilibrium strategy is given below in three cases specified by the market size Λ.*

a. ***Small market size.*** *If $\Lambda < \mu/(2-\alpha)$, the equilibrium strategy is*

$$(q_J^o, q_o^o) = \begin{cases} \mathbf{0}, & \text{if } \frac{1}{\mu} \ge \max\left\{\frac{V-P-C_h}{C}, \frac{V-P}{(2-\alpha)C}\right\}, \\ \left(\widetilde{q}_J^o, 0\right), & \text{if } \max\left\{\frac{1}{\mu}, \frac{1}{\delta}\right\} < \frac{V-P-C_h}{C} < \frac{1}{\mu-\Lambda}, \\ (1, 0), & \text{if } \frac{1}{\delta} < \frac{1}{\mu-\Lambda} \le \frac{V-P-C_h}{C}, \\ \left(1, \widehat{q}_o^o\right), & \text{if } \frac{1}{\mu-\Lambda} \le \frac{1}{\delta} \le \min\left\{\frac{1}{\mu-(2-\alpha)\Lambda}, \frac{(V-P)}{(2-\alpha)C}\right\}, \\ \left(\widehat{q}_J^o, 1\right), & \text{if } \frac{1}{\mu} < \frac{V-P}{(2-\alpha)C} < \min\left\{\frac{1}{\mu-(2-\alpha)\Lambda}, \frac{1}{\delta}\right\} \\ (1, 1), & \text{if } \frac{1}{\mu-(2-\alpha)\Lambda} \le \frac{V-P}{(2-\alpha)C} \ \& \ \frac{1}{\mu-(2-\alpha)\Lambda} < \frac{1}{\delta}. \end{cases} \tag{15}$$

b. ***Medium market size.*** *If $\mu/(2-\alpha) \le \Lambda < \mu$, the equilibrium strategy is*

$$(q_J^o, q_o^o) = \begin{cases} \mathbf{0}, & \text{if } \frac{1}{\mu} \ge \max\left\{\frac{V-P-C_h}{C}, \frac{V-P}{(2-\alpha)C}\right\}, \\ \left(\widetilde{q}_J^o, 0\right), & \text{if } \max\left\{\frac{1}{\mu}, \frac{1}{\delta}\right\} < \frac{V-P-C_h}{C} < \frac{1}{\mu-\Lambda}, \\ (1, 0), & \text{if } \frac{1}{\delta} < \frac{1}{\mu-\Lambda} \le \frac{V-P-C_h}{C}, \\ \left(1, \widehat{q}_o^o\right), & \text{if } \frac{1}{\mu-\Lambda} \le \frac{1}{\delta} \le \frac{(V-P)}{(2-\alpha)C}, \\ \left(\widehat{q}_J^o, 1\right), & \text{if } \frac{1}{\mu} < \frac{V-P}{(2-\alpha)C} < \frac{1}{\delta}. \end{cases} \tag{16}$$

c. ***Large market size.*** *If $\Lambda \ge \mu$, the equilibrium strategy is*

$$(q_J^o, q_o^o) = \begin{cases} \mathbf{0}, & \text{if } \frac{1}{\mu} \ge \max\left\{\frac{V-P-C_h}{C}, \frac{V-P}{(2-\alpha)C}\right\}, \\ \left(\widetilde{q}_J^o, 0\right), & \text{if } \max\left\{\frac{1}{\mu}, \frac{1}{\delta}\right\} < \frac{V-P-C_h}{C}, \\ \left(\widehat{q}_J^o, 1\right), & \text{if } \frac{1}{\mu} < \frac{V-P}{(2-\alpha)C} < \frac{1}{\delta}. \end{cases} \tag{17}$$

where $\widetilde{q}_J^o = \frac{\mu - \frac{C}{V-P-C_h}}{\Lambda}$, $\widehat{q}_o^o = \frac{\mu-\Lambda-\delta}{(1-\alpha)\Lambda}$, $\widehat{q}_J^o = \frac{\mu - \frac{(2-\alpha)C}{V-P}}{(2-\alpha)\Lambda}$ and $\delta = \frac{(1-\alpha)C}{C_h}$.

Please refer to Appendix for the proof.

4.3. Impact of service fee and expectation-meeting probability

In this subsection, we examine the impact of the service fee P and expectation-meeting probability α on customers' equilibrium strategy in both online-exchange and onsite-exchange models. We first study the effect of P.

Proposition 1 (*Impact of P*)**.** *In both online-exchange and onsite-exchange models, the equilibrium order-placing probability q_J is nonincreasing in P; the online-service probability q_o is nondecreasing in P.*

The monotonicity result in q_J is evident, because a higher price P discourages all customers from joining the system, regardless of the exchange policies. To understand the monotonicity in q_o, note that the lower congestion (due to a higher P) is able to achieve a bigger reduction in the delay cost for online customers than for onsite customers (because unsatisfied online customers will have to visit the service facility for a second time), thus incentivizing more customers to select the online service channel. We remark that Proposition 1 serves as a basis to obtain the optimal service fee in order to maximize the revenue (See Section 5.1 for details). We next give a numerical example to visualize results in Proposition 1. Fig. 2 confirms that the online-channel-selecting probability is non-decreasing in P under both exchange policies.

Proposition 2 (*Impact of α*)**.** *In both models, the online-channel-selecting probability q_o is nondecreasing in the expectation-meeting probability α; the equilibrium order-placing probability q_J is non-monotonic in α.*

Please refer to Appendix for the proof.

A bigger α makes the online service channel more trustworthy because fewer customers may have to revisit the store for exchange. Hence, the online-channel-selecting probability increases as α increases. On the other hand, the non-monotonicity of q_J is less straightforward: When α is sufficiently small, online customers are unlikely to receive a product that meets their expectations so that all customers (if joining) choose the onsite service channel, in which case α has no impact on customers' order-placing probability q_J. As α increases, customers are gradually leaning towards the online service channel. But such a change increases the overall customer delay (because online service is the source of product exchange) which impedes customers from joining the system (so that q_J is non-increasing), and in Fig. 3, we observe a sudden drop of q_J. When α is sufficiently close to 1, the highly reliable online service attracts more customers. Such an effect successfully reduces the overall queueing delay (because online service without future feedback has a smaller disutility than onsite service) and invites more customers to join the system.

Once again, we provide a numerical example to illustrate our findings in Proposition 2, see Fig. 3. Similar to results in Fig. 2, customers begin to switch from the onsite channel to online channel as α reaches a critical level.

4.4. Performance comparison: onsite exchange vs. online exchange

We next provide a performance comparison of the two product-exchange models; we aim to inform the service provider of the one that generates a higher revenue. In this subsection we restrict our attention to a fixed service fee P, so it suffices to compare the system throughput under the two policies: $TH^s = \Lambda q_J^s$ and $TH^o = \Lambda q_J^o$.

Proposition 3 (*Throughput Comparison with an Exogenous Service Fee*)**.** *We consider the following cases specified by the expectation-meeting probability α and the market size Λ.*

a. When $\alpha < \underline{\alpha}$, the two throughput functions are identical.
b. When $\alpha > \bar{\alpha}$, the online-exchange throughput is at least as high as the onsite-exchange throughput.
c. When $\alpha \in [\underline{\alpha}, \bar{\alpha}]$,
   - *if the market size $\Lambda \le \underline{\Lambda}$, the two throughput functions coincide;*
   - *if the market size $\Lambda > \underline{\Lambda}$, the online-exchange throughput is higher (lower) than the onsite-exchange throughput when price $P \ge \bar{P}$ ($P < \bar{P}$),*

where $\underline{\alpha} = \frac{V-P-2C_h}{V-P-C_h}$, $\bar{\alpha} = 1 - \frac{C_h}{V-P}$, $\underline{\Lambda} = \frac{\mu}{2-\alpha} - \frac{C}{V-P-(1-\alpha)C_h}$, $\bar{P} = V - C_h - \frac{\alpha C_h}{1-\alpha}$.

Please refer to Appendix for the proof.

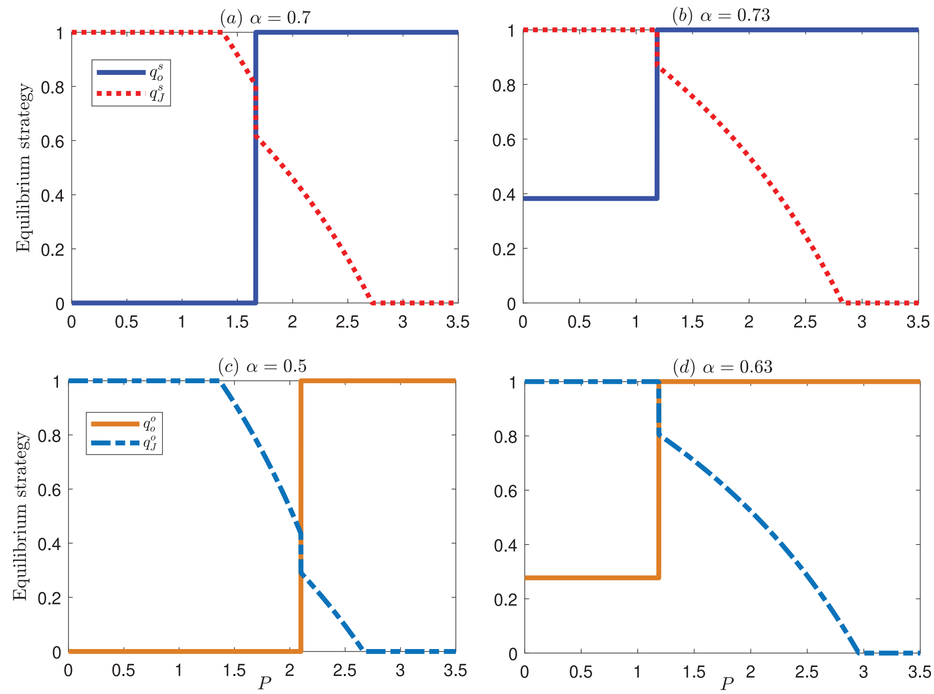


Fig. 2. The impact of price P on equilibrium strategy of the onsite-exchange model (top panel) and online-exchange model (bottom panel), with $V = 6$, $\mu = 0.9$, $\Lambda = 0.3$, $C = 2$, $C_h = 1.3$, $P \in [0, 3.5]$.

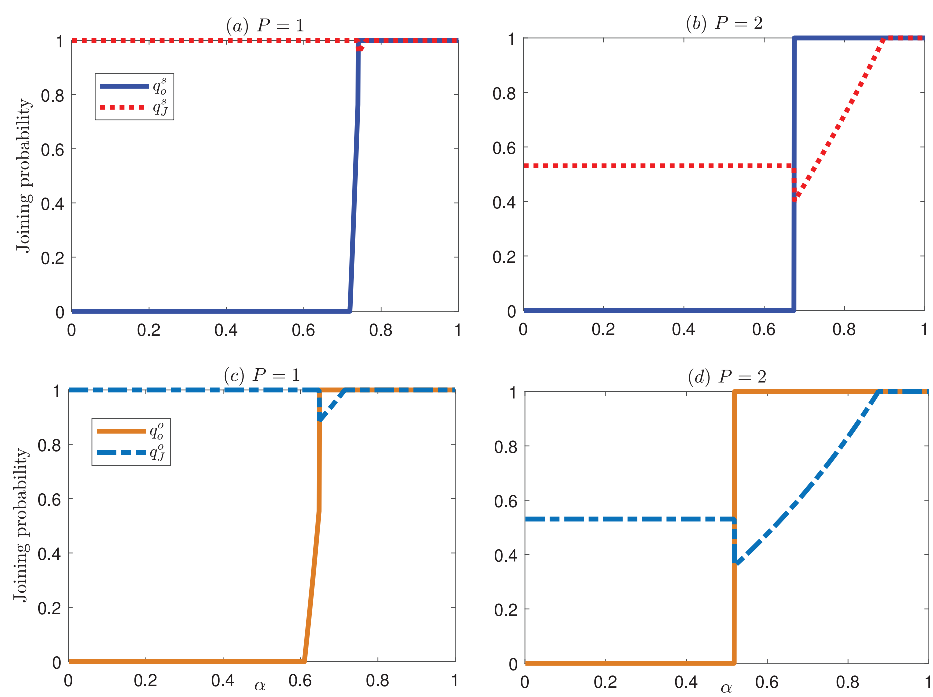


Fig. 3. The impact of expectation-meeting probability α on the equilibrium strategy of the onsite-exchange model (top panel) and online-exchange model (bottom panel), with $V = 6$, $\mu = 0.9$, $\Lambda = 0.3$, $C = 2$, $C_h = 1.3$, $P = 1, 2$, $\alpha \in (0, 1)$.

Remark 3 (*Onsite vs. Online*)**.** Here are the intuitions. When online service is unreliable (i.e., α is sufficiently small), nearly no one will select the online channel, so that the specific exchange policy (online or onsite) has no impact on the system performance. On the other hand, if online service is highly reliable (α is large), almost all customers prefer to order online regardless which exchange policy is being implemented. In this case, because online-exchange policy waives the inconvenience cost for exchanging a product, it attracts more customers to place orders, thus inducing a bigger throughput than the onsite-exchange policy.

Finally, we consider the case α is intermediate (i.e., $\alpha \in (\underline{\alpha}, \bar{\alpha})$ for some $\underline{\alpha} < \bar{\alpha}$). If Λ is small (i.e., $\Lambda < \underline{\Lambda}$ for some $\underline{\Lambda}$), all potential customers join the system, inducing the identical system throughput for the two models. See Fig. 4 for an example. When Λ is large, the system's congestion level is highly dependent on the service price P: A smaller P leads to a higher system congestion level (a longer waiting

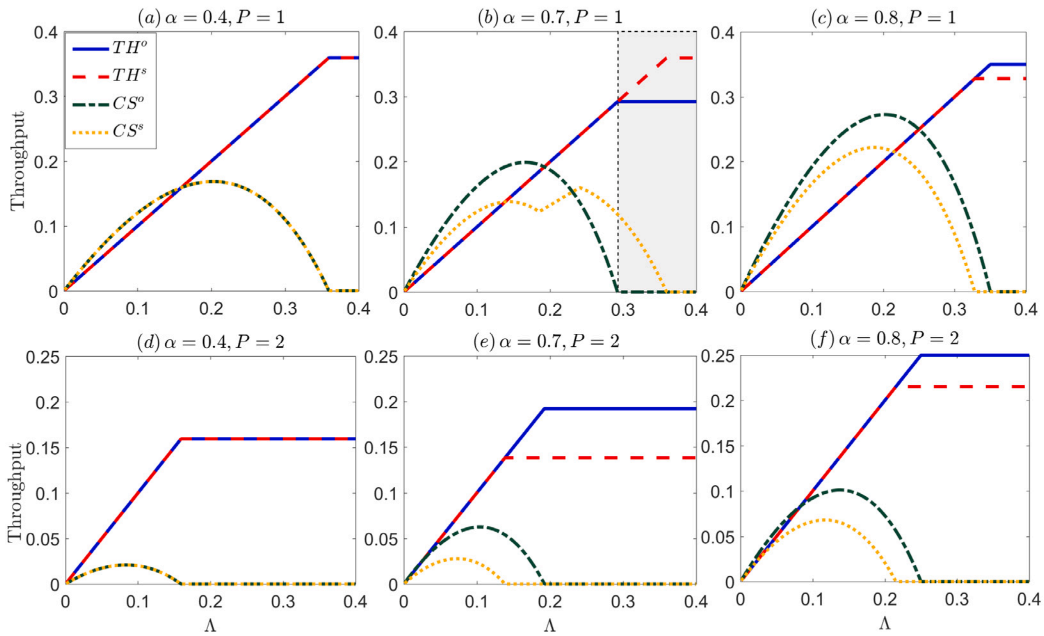


Fig. 4. Comparison of throughput and consumer surplus in the two exchange models, with $V = 6$, $\mu = 0.9$, $C = 2$, $C_h = 1.3$, $0 \leq \Lambda \leq 0.4$, $P = 1, 2$, $\alpha = 0.4, 0.7, 0.8$.

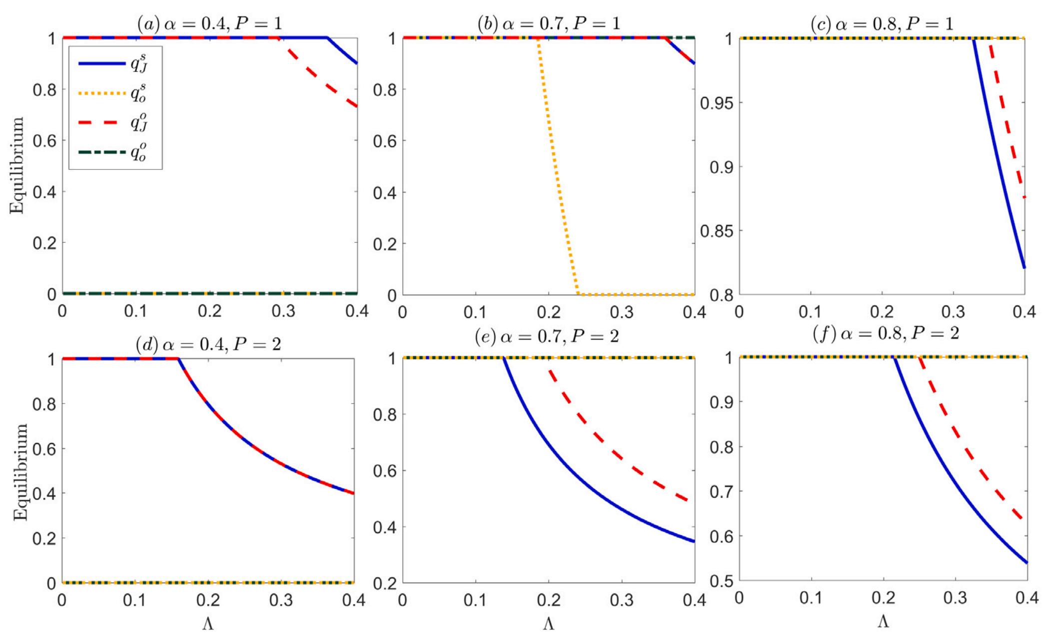


Fig. 5. Comparison of equilibrium strategies in the two exchange models, with $V = 6$, $\mu = 0.9$, $C = 2$, $C_h = 1.3$, $0 \leq \Lambda \leq 0.4$, $P = 1, 2$, $\alpha = 0.4, 0.7, 0.8$.

time), which diverts more customers from online to onsite channel (see our previous discussion in Proposition 1). This effect is more pronounced in the onsite-exchange model because online customers receiving dissatisfying products will have to experience the queueing process again. Therefore, a lower average delay in the onsite-exchange model is able to attract more customers to join for service. When the price P is high, more customers choose the online service channel, and in this case the online-exchange model yields a higher throughput because it avoids the inconvenience cost C_h completely.

Below we consider a numerical example to visualize results in Proposition 3. Besides the system throughput, we also consider the *consumer surplus* (CS), which is the sum of all customers' utility gain. Specifically, the CS under two models are defined as

$$CS^s = \Lambda q_J^s[q_o^s U_o^s + (1-q_o^s)U_s^s] \quad \text{and} \quad CS^o = \Lambda q_J^o[q_o^o U_o^o + (1-q_o^o)U_s^o], \tag{18}$$

where $U_o^s, U_s^s, U_o^o, U_s^o$ are defined in Eqs. (4)–(5), (12)–(13) respectively. In Fig. 4 we plot the throughput and consumer surplus (the sum of utilities of all consumers) of both online-exchange and onsite-exchange models as functions of Λ, P and α. Observations in Fig. 4 on the throughput are consistent with findings in Proposition 3. However, an interesting case for CS is when α is intermediate and P is small (panel (b)), where onsite exchange is more effective in both the throughput and CS when Λ is large (see the shaded area): As Λ increases, the online-exchange model is the first to reach its throughput saturation point, yielding a constant throughput and zero customer surplus, whereas the onsite-exchange model continues to improve its throughput and render a positive customer surplus. See Fig. 5 for the detailed equilibrium strategies.

On the customer utility. We next investigate how the market size impacts the customer utility functions. In Fig. 6, and we plot customer

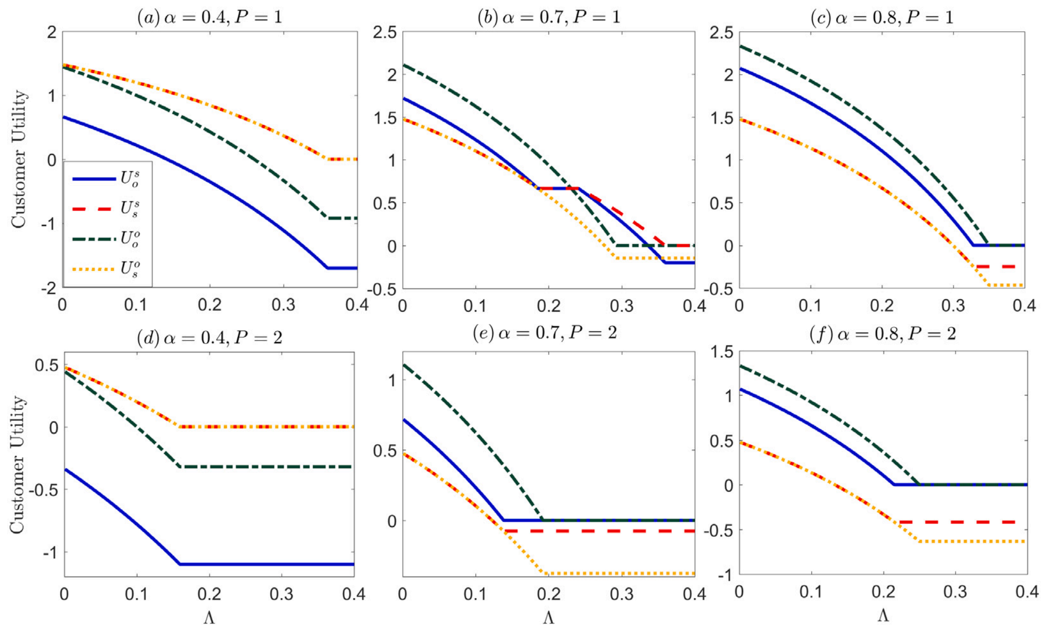


Fig. 6. Comparison of customer utility in the two exchange models, with $V = 6$, $\mu = 0.9$, $C = 2$, $C_h = 1.3$, $0 \leq \Lambda \leq 0.4$, $P = 1, 2$, $\alpha = 0.4, 0.7, 0.8$.

utility under both service channels in two models. First, as the market size increases, all customer utility functions decrease.

Next, when α is sufficiently small (e.g., $\alpha = 0.4$ as in panel (a)), in both exchange models, customers' utility function of the onsite channel are higher than those of the online channel ($U_s^s > U_o^s; U_s^o > U_o^o$), so that all customers select the onsite service channel. Moreover, the structure of customer utility functions in the onsite channel are identical under both models because no exchange will be needed (i.e., $U_s^s = U_s^o$). When α is sufficiently large (e.g., $\alpha = 0.8$ as in panel (c)), in both exchange models, customers' utility functions of online channel are higher than those of the onsite channel (*i.e.*, $U_o^s > U_s^s; U_o^o > U_s^o$), so all customers select the online channel. Indeed, the online-exchange model prevails because its high satisfactory probability helps reduce customers' delay costs.

The case of the intermediate α (e.g., $\alpha = 0.7$ as in panel (b)) reveals some interesting observations: here the ordering of the utility functions depends on the market size Λ and price P. If the market size is sufficiently small (e.g., $\Lambda < 0.1$), all potential customers join the system and select the online service channel. This is because that the online channel helps reduce the inconvenience cost. As the market size increases, customers' utility will be dependent on the service price P: When the price P is high, we observe that $U_o^s \geq U_s^s$ and $U_o^o \geq U_s^o$. A high service fee can help reduce the system's waiting time, making the online service channel more appealing. If the price P is small, the system's waiting time is larger. In the online exchange model, the online service channel becomes more advantageous because it helps reduce the inconvenience cost; while in the onsite exchange model, the onsite service channel is more popular because it can reduce the additional waiting costs during the exchange process.

5. Extensions

We study two extensions of our base model. First, in Section 5.1 we allow the service fee to be a decision variable which can be used to further improve the revenue. Next, in Section 5.2 we allow the inconvenience cost to be a random variable to capture customers' heterogeneity.

5.1. Endogenous price

Previously we have examined customers' equilibrium strategy and the system throughput in the two product-exchange models with the service fee held fixed. This analysis is relevant to many practices where the service provider is unable to adjust the price of the product (e.g., when retail prices are value-based or competition-based; see, for example, [36]). Nevertheless, pricing is not only feasible but also crucial in many other practices (e.g., on-demand platforms such as [37]).

In this subsection, we consider the setting of *endogenous price*, where the retailer is allowed to adjust the price P to maximize the revenue. We study the corresponding customer responses and the performance under the two exchange policies. Denote by $\Pi(P)$ the revenue per unit time for the retailer under price P, the revenue maximization problem is

$$\max_{P \geq 0} \Pi(P) = \Lambda_e(P)P, \tag{19}$$

where $\Lambda_e(P)$ is the effective external arrival rate given P.

In the rest of this subsection, we focus on the small-market case with $\mu > (2-\alpha)\Lambda$ (Case (a) in Theorem 1). We remark that the analysis for the other two cases are similar and can be conducted following an identical road map, so we omit them to avoid tediousness.

5.1.1. Onsite exchange model with endogenous service fee

We first consider the onsite-exchange model (we add subscript "s" to the optimal price P^*). To characterize customers' equilibrium strategy, we define the following constants which, as we will see soon, are candidate pricing solutions.

$$\widetilde{P}_s = V - C_h - \sqrt{\frac{C(V - C_h)}{\mu}}, \quad \hat{P}_s = \bar{V} - \sqrt{\frac{(2-\alpha)C\bar{V}}{\mu}}. \tag{20}$$

In what follows, we provide results that can be used to obtain the retailer's optimal price P_s^*, where the subscript "s" indicates the onsite exchange policy. Let $x \vee y \equiv \max(x, y)$ and $x \wedge y \equiv \min(x, y)$. Due to the complex nature of the problem, a closed-form expression of P_s^* is difficult. Hence, we provide a partial description by computing the optimal prices assuming a specific structure of customers' equilibrium strategy (q_J^s, q_o^s).

Theorem 3 (*Onsite Exchange with Endogenous Pricing*)**.** *In the onsite-exchange model, the service provider's optimal price can be described in the following three cases specified by customers' equilibrium strategy* (q_J^s, q_o^s).

i. For equilibrium strategy $(q_J^s, q_o^s) = (\widetilde{q}_J^s, 0)$, *(i.a) If* $\mu - \frac{C}{V-C_h} < 0$, $P_s^* = V - C_h - \frac{C}{\mu-\Lambda}$;

(i.b) If $\mu - \frac{C}{V-C_h} \geq 0$ *and* $V - C_h - \frac{C}{\mu} \geq V - C_h - \frac{C}{\eta}$,

$$P_s^* = \left(V - C_h - \frac{C}{\mu - \Lambda}\right) \vee \widetilde{P}_s \wedge \left(V - C_h - \frac{C}{\eta}\right);$$

(i.c) If $\mu - \frac{C}{V-C_h} \geq 0$ *and* $V - C_h - \frac{C}{\mu} < V - C_h - \frac{C}{\eta}$,

$$P_s^* = \left(V - C_h - \frac{C}{\mu - \Lambda}\right) \vee \widetilde{P}_s \wedge \left(V - C_h - \frac{C}{\mu}\right).$$

ii. *For equilibrium strategy* $(q_J^s, q_o^s) = (1, \hat{q}_o^s)$,

$$P_s^* = V - \frac{C}{\eta\alpha};$$

iii. *For equilibrium strategy* $(q_J^s, q_o^s) = (\hat{q}_J^s, 1)$,

(iii.a) If $\mu - \frac{(2-\alpha)C}{V-(1-\alpha)C_h} < 0$,

$$P_s^* = \max\left\{\bar{V} - \frac{(2-\alpha)C}{\eta}, \bar{V} - \frac{(2-\alpha)C}{\mu - (2-\alpha)\Lambda}\right\};$$

(iii.b) If $\mu - \frac{(2-\alpha)C}{V-(1-\alpha)C_h} \geq 0$ *and* $\frac{(2-\alpha)C}{\eta} \leq \frac{(2-\alpha)C}{\mu-(2-\alpha)\Lambda}$,

$$P_s^* = \left(\bar{V} - \frac{(2-\alpha)C}{\eta}\right) \vee \hat{P}_s \wedge \left(\bar{V} - \frac{(2-\alpha)C}{\mu}\right);$$

(iii.c) If $\mu - \frac{(2-\alpha)C}{V-(1-\alpha)C_h} \geq 0$ *and* $\frac{(2-\alpha)C}{\eta} > \frac{(2-\alpha)C}{\mu-(2-\alpha)\Lambda}$,

$$P_s^* = \left(\bar{V} - \frac{(2-\alpha)C}{\mu - (2-\alpha)\Lambda}\right) \vee \hat{P}_s \wedge \left(\bar{V} - \frac{(2-\alpha)C}{\mu}\right).$$

Please refer to Appendix for the proof.

5.1.2. Online exchange model with endogenous service fee

We next study the online exchange model (we add subscript "o" to P^*). To characterize customers' equilibrium strategy, we define the following constants which are candidate pricing solutions.

$$\widetilde{P}_o = V - C_h - \sqrt{\frac{C(V - C_h)}{\mu}} \quad \text{and} \quad \hat{P}_o = V - \sqrt{\frac{(2-\alpha)CV}{\mu}}. \tag{21}$$

Similar to Theorem 3, we next provide a partial description of the optimal price in accordance with every case of the equilibrium strategy.

Theorem 4 (*Online Exchange with Endogenous Pricing*)**.** *In the online-exchange model, the service provider's optimal price can be described in the following three cases specified by customers' equilibrium strategy* (q_J^o, q_o^o).

i. *For equilibrium strategy* $(q_J^o, q_o^o) = (\widetilde{q}_J^o, 0)$,

(i.a) if $\mu - \frac{C}{V-C_h} < 0$, $P_o^* = V - C_h - \frac{C}{\mu-\Lambda}$;

(i.b) if $\mu - \frac{C}{V-C_h} \geq 0$ *and* $V - C_h - \frac{C}{\mu} \geq V - C_h - \frac{C}{\delta}$,

$$P_o^* = \left(V - C_h - \frac{C}{\mu - \Lambda}\right) \vee \widetilde{P}_o \wedge \left(V - C_h - \frac{C}{\delta}\right);$$

(i.c) if $\mu - \frac{C}{V-C_h} \geq 0$ *and* $V - C_h - \frac{C}{\mu} < V - C_h - \frac{C}{\delta}$,

$$P_o^* = \left(V - C_h - \frac{C}{\mu - \Lambda}\right) \vee \widetilde{P}_o \wedge \left(V - C_h - \frac{C}{\mu}\right).$$

ii. *For equilibrium strategy* $(q_J^o, q_o^o) = (1, \hat{q}_o^o)$,

$$P_o^* = V - \frac{(2-\alpha)C}{\delta};$$

iii. *For equilibrium strategy* $(q_J^o, q_o^o) = (\hat{q}_J^o, 1)$,

(iii.a) if $\mu - \frac{(2-\alpha)C}{V} < 0$,

$$P_o^* = \max\left\{V - \frac{(2-\alpha)C}{\delta}, V - \frac{(2-\alpha)C}{\mu - (2-\alpha)\Lambda}\right\};$$

(iii.b) If $\mu - \frac{(2-\alpha)C}{V} \geq 0$ *and* $\frac{(2-\alpha)C}{\delta} \leq \frac{(2-\alpha)C}{\mu-(2-\alpha)\Lambda}$,

$$P_o^* = \left(V - \frac{(2-\alpha)C}{\delta}\right) \vee \hat{P}_o \wedge \left(V - \frac{(2-\alpha)C}{\mu}\right);$$

(iii.c) If $\mu - \frac{(2-\alpha)C}{V} \geq 0$ *and* $\frac{(2-\alpha)C}{\delta} > \frac{(2-\alpha)C}{\mu-(2-\alpha)\Lambda}$,

$$P_o^* = \left(V - \frac{(2-\alpha)C}{\mu - (2-\alpha)\Lambda}\right) \vee \hat{P}_o \wedge \left(V - \frac{(2-\alpha)C}{\mu}\right).$$

Please refer to Appendix for the proof.

5.1.3. Revenue comparison with endogenous service fee

In this subsection, we compare the revenue of the two product-exchange models when the service provider adopts the optimal price according to (19). We pay especial attention to the scenarios specified by the market size Λ and the expectation-meeting probability α.

Proposition 4 (*Revenue Under Endogenous Pricing*)**.** *The optimal prices in both product-exchange models are non-increasing in* Λ. *To compare the optimal revenue, we consider three cases specified by the expectation-meeting probability* α.

a. *When* $\alpha < \underline{\alpha}_1$, *revenue of the two models are identical.*
b. *When* $\alpha > \bar{\alpha}_1$, *the online-exchange model generates a higher revenue than the onsite-exchange model.*
c. *When* $\alpha \in [\underline{\alpha}_1, \bar{\alpha}_1]$, *the online-exchange model yields a higher (lower) revenue than the onsite-exchange model when the market size is lower (higher) than* $\underline{\Lambda}'$,

where $\underline{\alpha}_1$ *is uniquely solved by* $C_h^2\mu = (2-\alpha)(1-\alpha)^2 C(V - C_h)$, *and* $\bar{\alpha}_1 = \frac{\sqrt{\mu C(V-C_h)}}{\mu C_h + \sqrt{\mu C(V-C_h)}}$ *and* $\underline{\Lambda}' = \frac{\mu}{2-\alpha} - \frac{\mu C}{\alpha\mu C_h + \sqrt{\mu C(V-C_h)}}$.

Please refer to Appendix for the proof.

We visualize our findings by considering an example to numerically compare the optimal revenue and its corresponding price under two policies. In Fig. 7, we plot the optimal revenue Π_s^* and Π_o^* (bottom panels) and the optimal price P_s^* and P_o^* (top panels) for $0 \leq \Lambda \leq 0.4$ with $\alpha = 0.4, 0.7, 0.8$.

Fig. 7 gives the following implications.

i. When online service is highly unreliable (e.g., $\alpha = 0.4$), all joining customers choose the onsite service channel, which yields no feedback orders at all. Hence, the specific exchange policy plays no role, and the two revenue curves coincide.
ii. On the other hand, when online service is sufficiently reliable (e.g., $\alpha = 0.8$), all joining customers choose the online service channel. Because the online-exchange model induces a lower total cost for joining customers (it spares the cost C_h), more customers are willing to join. This gives the service provider an opportunity to achieve higher revenue by increasing the service fee.
iii. The most sophisticated is the middle ground case with an intermediate α, see the middle plots in Fig. 7. In this case, the online-exchange model yields a higher (lower) revenue when the mark size Λ is small (large). To gain insights into this behavior, note that increasing Λ will divert a bigger customer flow from the online channel to onsite channel (see our previous discussion following Proposition 1). Although this happens to both models, this effect is more significant in the onsite-exchange model due to the extra inconvenience cost. This behavior effectively mitigates the system congestion which in turn allows more customers to join the system, hence a higher revenue.

When the service fee is endogenous (under the setting of the present section), the performance ranking of the two models remains nearly unchanged when α is small (see panel (a) in Figs. 4 and 7). When α is large, optimal pricing allows the online-exchange model to establish a bigger advantage (see panel (c) in Figs. 7), because the service provider may increase the price to fetch off more consumer surplus. When α is

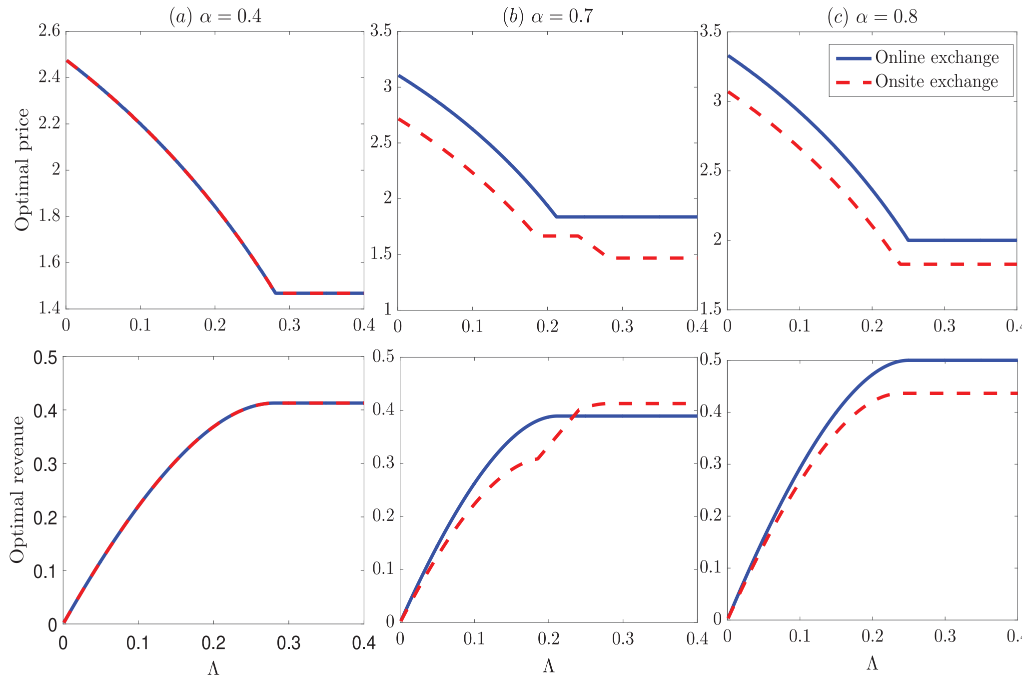


Fig. 7. Comparison of the revenue and price in the two product-exchange models, with $V = 6$, $\mu = 0.9$, $C = 2$, $C_h = 1.3$.

medium, we know that, according to Fig. 4, online exchange gives a higher (lower) revenue when the price P is high (low). Because the optimal price P^* decreases in Λ, online exchange ought to outperform onsite exchange when Λ is small (i.e., P^* is high), but its dominance on the revenue should be overturned as the market size Λ (optimal price P^*) becomes sufficiently large (low); also see Remark 3.

Finally, we study the impact of the expectation-meeting probability α on the optimal revenue in the two models. We do so via a numerical example, see Fig. 8. Consistent with our earlier findings under the setting of a fixed P, Fig. 8 confirms that the revenue is not always increasing in α (see Fig. 3 for example). Indeed, when α is medium, increasing α will drive more customers to join the online service channel (q_o increases in α) and thus prolong the overall customer waiting time (because product exchange is generated from online service only) which impedes customers from joining the system and yields a lower revenue. This explains why all revenue curves have a little "dip". But as α continues to grow, all joining customers will eventually choose the online channel, and because the online-exchange policy eliminates the inconvenience cost C_h, the effective arrival rate increases so that the service provider is able to harvest higher revenue.

5.2. Heterogeneous inconvenience cost

Our analysis in previous sections assumes a constant inconvenience cost for exchange customers. However, such an inconvenience cost may vary from customer to customer in practice, we now consider an extension to include a *heterogeneous* inconvenience cost. Specifically, we now assume that C_h is a random variable which is uniformly distributed in $[0, C_H]$ for some $C_H > 0$.

5.2.1. Onsite exchange model

Because the inconvenience cost is now a heterogeneous factor, the utility function will depend on the specific customer's inconvenience cost. As we will see later, the equilibrium strategy will exhibit threshold structure in the value of C_h. In the onsite exchange model, given an inconvenience cost $C_h \in [0, C_H]$, the expected utilities of an online customer and an onsite customer are

$$U_o^s(p_s, p_o; C_h) = V - P - (1-\alpha)C_h - (2-\alpha)CW(p_s, p_o), \tag{22}$$

$$U_s^s(p_s, p_o; C_h) = V - P - C_h - CW(p_s, p_o), \tag{23}$$

where p_s and p_o are the fractions of customers who order onsite and online, with $p_s + p_o \leq 1$, and $W(p_s, p_o) = \frac{1}{\mu - \Lambda p_s - (2-\alpha)\Lambda p_o}$ denotes the customers' expected waiting time under p_s and p_o. Because both utility functions (22) and (23) are decreasing in C_h, for any given (p_s, p_o), there exists unique solutions to equations $U_o^s(p_s, p_o; C_h) = 0$ and $U_s^s(p_s, p_o; C_h) = 0$, which we refer to as $\hat{C}_h$ and $\tilde{C}_h$ respectively. In addition, there exists a unique solution to the equation $U_o^s(p_s, p_o; C_h) = U_s^s(p_s, p_o; C_h)$, which we denote as $\bar{C}_h$. Specifically, we have

$$\tilde{C}_h = V - P - CW(p_s, p_o), \quad \hat{C}_h = \frac{V - P - (2-\alpha)CW(p_s, p_o)}{1-\alpha},$$
$$\bar{C}_h = \frac{(1-\alpha)CW(p_s, p_o)}{\alpha}.$$

We next give customers equilibrium using the above-introduced constants.

Theorem 5 (*Onsite Exchange Model with Heterogeneous Inconvenience Cost*). *Under the onsite exchange policy, we give the equilibrium strategy as follows.*

$$(p_s^e, p_o^e) = \begin{cases} (1,0), & \text{if } \frac{\alpha C_H}{(1-\alpha)C} < W(1,0) \leq \frac{V-P-C_H}{C}, \\ \left(p_s^1, 1-p_s^1\right), & \text{if } W(p_s^1, 1-p_s^1) \leq \min\left\{\frac{\alpha(V-P)}{C}, \frac{\alpha C_H}{(1-\alpha)C}, \frac{(V-P)-(1-\alpha)C_H}{(2-\alpha)C}\right\}, \\ (p_s^2, p_o^2), & \text{if } \frac{(V-P)-(1-\alpha)C_H}{(2-\alpha)C} < W(p_s^2, p_o^2) \leq \frac{\alpha(V-P)}{C}, \\ \left(p_s^3, 0\right), & \text{if } W(p_s^3, 0) > \max\left\{\frac{\alpha(V-P)}{C}, \frac{V-P-C_H}{C}\right\}. \end{cases} \tag{24}$$

where p_s^1, p_s^3, p_s^2 and p_o^2 solve equations $\alpha p_s^1 C_H = (1-\alpha)CW(p_s^1, 1-p_s^1)$, $p_s^3 C_H = V - P - CW(p_s^3, 0)$, $\alpha p_s^2 C_H = (1-\alpha)CW(p_s^2, p_o^2)$, and

$$p_o^2 C_H = \frac{V - P - (2-\alpha)CW(p_s^2, p_o^2)}{1-\alpha} - \frac{(1-\alpha)CW(p_s^2, p_o^2)}{\alpha}.$$

Please refer to Appendix for the proof.

5.2.2. Online exchange model

In the online exchange model, given an inconvenience cost $C_h \in [0, C_H]$, the expected utilities of an online customer and an onsite

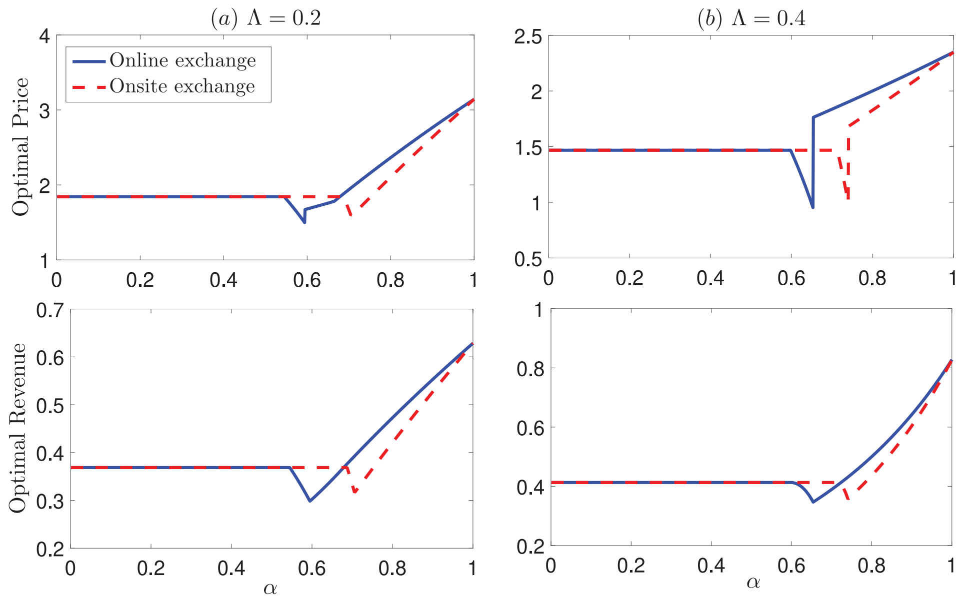


Fig. 8. Comparison of the optimal price and revenue in the two product-exchange models, with $V = 6$; $\mu = 0.9$; $C = 2$; $C_h = 1.3$.

customer are

$$U_o^o(p_s, p_o) = V - P - (2-\alpha)CW(p_s, p_o), \tag{25}$$

$$U_s^o(p_s, p_o; C_h) = V - P - C_h - CW(p_s, p_o). \tag{26}$$

Note that in (25) we omit the argument C_h because the utility function is independent of C_h. For a given (p_s, p_o), the utility function in (26) is decreasing in C_h. Hence, there exists a unique solution to the equation $U_s^o(p_s, p_o; C_h) = 0$, which we refer to as $\underline{C}_h$. In addition, there exists a unique solution to the equation $U_o^o(p_s, p_o) = U_s^o(p_s, p_o; C_h)$, in which we denote as C_h'. Specifically, we have

$$\underline{C}_h = V - P - CW(p_s, p_o), \qquad C_h' = (1-\alpha)CW(p_s, p_o).$$

Theorem 6 (*Onsite Exchange Model with Heterogeneous Inconvenience Cost*)**.** *Under the onsite exchange policy, we give the equilibrium strategy as follows.*

$$(p_s^e, p_o^e) = \begin{cases} (1,0), & \text{if} \quad \frac{C_H}{(1-\alpha)C} < W(1,0) \le \frac{V-P-C_H}{C}, \\ \left(p_s^1, 1-p_s^1\right), & \text{if} \quad W(p_s^1, 1-p_s^1) \le \min\left\{\frac{V-P}{(2-\alpha)C}, \frac{C_H}{(1-\alpha)C}\right\}, \\ \left(p_s^3, 0\right), & \text{if} \quad W(p_s^3, 0) > \max\left\{\frac{V-P}{(2-\alpha)C}, \frac{V-P-C_H}{C}\right\}. \end{cases} \tag{27}$$

where p_s^1 and p_s^3 solve equations $p_s^1 C_H = (1-\alpha)CW(p_1^s, 1-p_s^1)$ and $p_s^3 C_H = V - P - CW(p_s^3, 0)$.

Please refer to Appendix for the proof.

6. Conclusion

Motivated by the rapid growth of online retailing, we study an omnichannel service queueing model. The omnichannel aspect of the model allows customers to either place an online order (e.g., on a computer or mobile app), or place an onsite order by physically visiting the service facility. Although the online service channel has many appeals, it gives rise to undesired feedback orders requesting return and exchanges; on the other hand, the physical travel to a service facility can be inconvenience and costly, but it gives customers sufficient opportunities to interact with the product, thus can largely avert feedback orders. In this paper, we consider two practical exchange policies: online exchange (e.g., by mail), and onsite exchange (e.g., in-store dropoff). We carefully examine how these two exchange policies

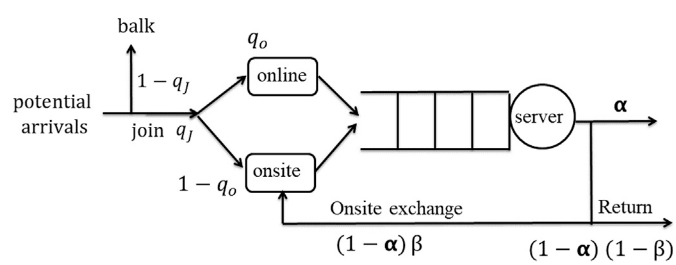


Fig. 9. An omnichannel retailing queue with product exchange and return.

shape customers' behavior; we aim to inform the service provider of the optimal exchange policy with the objective of maximizing the system's total revenue. We study two settings: the service fee is exogenous (not a decision variable), and the service fee is endogenous (part of the service provider's decision). Our results show that the online exchange policy is a double-edged sword: On the one hand, it helps eliminate the inconvenience cost for exchange customers to revisit the store; on the other hand, it can trigger more feedback orders and render a higher system congestion level, which in turn, deters future customers from placing orders. Specifically, we discovery that online exchange becomes an inferior policy (relative to onsite exchange) when the market size is large.

Limitations and future directions. There are several venues for future research. First, it will be interesting to consider the more general setting where online exchange also incurs an inconvenience cost C_o and carefully study how C_o compares to C_h. A second future direction is to consider endogenous product-exchange behavior, that is, a customer receiving a dissatisfying product from the online service channel may or may not pursue a product exchange. We envision that this extension will largely complicate the analysis of the customer equilibrium strategy which will now become a three-dimensional vector because customers make decisions in three steps (the two-dimensional decision in the present setting is already quite involved so this future work will require some serious efforts). Another potential generalization is to consider the case which allows customers to request for returning the product for a complete refund when they are unsatisfied with the product, see Fig. 9.

CRediT authorship contribution statement

Ke Sun: Conceptualization, Formal analysis, Methodology, Visualization, Original draft. **Yunan Liu:** Conceptualization, Writing – review & editing. **Xiang Li:** Supervision, Writing – review & editing.

Declaration of competing interest

There is NO conflict of interest.

Data availability

No data was used for the research described in the article.

Acknowledgments

This work was supported by Beijing Natural Science Foundation, China (No. 9244031), and National Natural Science Foundation of China (No. 71931001), and the Funds for First-class Discipline Construction, China (XK1802-5) and the Fundamental Research Funds for the Central Universities, China (ZY2338).

Appendix. Proofs

Proofs of Lemma 1. According to the customer utilities in the two service channels (4)–(5), we define the difference of these two functions as

$$\Delta^s(q_J,q_o) \equiv U_o^s(q_J,q_o) - U_s^s(q_J,q_o) = \alpha C_h - \frac{(1-\alpha)C}{\mu - \Lambda q_J[1+(1-\alpha)q_o]}. \quad \text{(A.1)}$$

For a given q_J, $\Delta^s(q_J,q_o)$ is decreasing in $q_o \in [0,1]$, and its maximum and minimum values are

$$\Delta^{\uparrow}(q_J) \equiv \max_{q_o} \Delta^s(q_J,q_o) = \Delta^s(q_J,0) = \alpha C_h - (1-\alpha)\frac{C}{\mu - \Lambda q_J},$$

$$\Delta^{\downarrow}(q_J) \equiv \min_{q_o} \Delta^s(q_J,q_o) = \Delta^s(q_J,1) = \alpha C_h - (1-\alpha)\frac{C}{\mu - \Lambda q_J[2-\alpha]}.$$

Let $\eta \equiv \frac{(1-\alpha)C}{\alpha C_h}$, we consider three cases specified by η:

(1) When $\frac{1}{\eta} < \frac{1}{\mu-\Lambda q_J}$, we have $\Delta^{\uparrow}(q_J) < 0$ so that $\Delta^s(q_J,q_o) < 0$ for all $q_o \in [0,1]$. Hence, the best response of a joining customer is to order onsite, yielding the equilibrium $q_o^s(q_J) = 0$.

(2) When $\frac{1}{\mu-\Lambda q_J} \le \frac{1}{\eta} \le \frac{1}{\mu-\Lambda q_J(2-\alpha)}$, there exists a unique solution q_o^* to the equation $\Delta^s(q_J,q_o^*) = 0$. In addition, we have $\Delta^s(q_J,q_o) > 0$ for $q_o \in [0,q_o^*)$ and $\Delta^s(q_J,q_o) < 0$ for $q_o \in (q_o^*,1]$. So the equilibrium order-selecting probability $q_o^s(q_J) = q_o^* = \frac{1}{1-\alpha}\left(\frac{\mu-\eta}{\Lambda q_J} - 1\right)$.

(3) When $\frac{1}{\eta} > \frac{1}{\mu-\Lambda q_J(2-\alpha)}$, we have $\Delta^{\downarrow}(q_J) > 0$ so that $\Delta^s(q_J,q_o) > 0$ for all $q_o \in [0,1]$. Hence, the best response for a joining customer is to order online, yielding the equilibrium $q_o^s(q_J) = 1$.

To show that the equilibrium channel-selecting probability $q_o^s(q_J)$ is weakly decreasing in q_J, we pick two probabilities $q_J^1 < q_J^2$ and denote the corresponding equilibrium channel-selecting probabilities as $q_o^{s,1}(q_J^1)$ and $q_o^{s,2}(q_J^2)$, respectively. Hence, we have that $x_1 \equiv \frac{1}{\mu-\Lambda q_J^1} < \frac{1}{\mu-\Lambda q_J^2} \equiv x_2$ and $y_1 \equiv \frac{1}{\mu-(2-\alpha)\Lambda q_J^1} < \frac{1}{\mu-(2-\alpha)\Lambda q_J^2} \equiv y_2$. In addition, we also have that $z_1 \equiv \frac{1}{1-\alpha}\left(\frac{\mu-\eta}{\Lambda q_J^1} - 1\right) > \frac{1}{1-\alpha}\left(\frac{\mu-\eta}{\Lambda q_J^2} - 1\right) \equiv z_2$. We next separately treat the following two cases:

(1) When $x_1 < x_2 \le y_1 < y_2$: According to the equilibrium results in Eq. (6), if $\frac{1}{\eta} \le x_2$, then $q_o^{s,1}(q_J^1) \ge q_o^{s,2}(q_J^2) = 0$; if $x_2 < \frac{1}{\eta} < y_1$, $q_o^{s,1}(q_J^1) = z_1 > z_2 = q_o^{s,2}(q_J^2)$; if $\frac{1}{\eta} \ge y_1$, $q_o^{s,1}(q_J^1) = 1 \ge q_o^{s,2}(q_J^2)$.

(2) When $x_1 < y_1 < x_2 < y_2$: If $\frac{1}{\eta} \le x_2$, then $q_o^{s,1}(q_J^1) \ge q_o^{s,2}(q_J^2) = 0$; If $\frac{1}{\eta} > x_2$, $q_o^{s,1}(q_J^1) = 1 \ge q_o^{s,2}(q_J^2)$.

In summary, we have $q_o^{s,1}(q_J^1) \ge q_o^{s,2}(q_J^2)$. □

Proof of Theorem 1. In Case (a) (i.e., the market size $\Lambda < \mu/(2-\alpha)$), there are three sub-cases:

(a.i) If $q_o^s = 0$, recall that the first case in Lemma 1 requires that $\frac{1}{\eta} < \frac{1}{\mu-\Lambda q_J}$, thus the order-placing probability satisfies the condition

$$q_J > \frac{\mu-\eta}{\Lambda}. \quad \text{(A.2)}$$

Subsequently, a joining customer's expected utility satisfies $U_J^s(q_J,0) = U_s^s(q_J,0) = V - P - C_h - \frac{C}{\mu-\Lambda q_J}$, where $U_J^s(q_J,0)$ is decreasing in $q_J \in [0,1]$.

(1) When $\frac{V-P-C_h}{C} \le \frac{1}{\mu}$, then $U_J^s(q_J,0) \le 0$ for all $q_J \in [0,1]$, so the best response of all arriving customers is to balk, which is an equilibrium if $q_J = 0 > \frac{\mu-\eta}{\Lambda}$, or equivalently, $\frac{1}{\mu} > \frac{1}{\eta}$.

(2) When $\frac{1}{\mu} < \frac{V-P-C_h}{C} < \frac{1}{\mu-\Lambda}$, there exists a unique solution $\widetilde{q}_J^s$ to the equation $U_J^s(q_J,0) = 0$, which is an equilibrium if $\widetilde{q}_J^s = \frac{\mu - \frac{C}{V-P-C_h}}{\Lambda} > \frac{\mu-\eta}{\Lambda}$ or equivalently, $\frac{V-P-C_h}{C} > \frac{1}{\eta}$.

(3) When $\frac{V-P-C_h}{C} \ge \frac{1}{\mu-\Lambda}$, then $U_J^s(q_J,0) \ge 0$ for all $q_J \in [0,1]$, so that joining becomes the best response for all customers which is an equilibrium if $1 > \frac{\mu-\eta}{\Lambda}$ or equivalently, $\frac{1}{\eta} < \frac{1}{\mu-\Lambda}$.

(a.ii) If $q_o^s \in (0,1)$, then customers' order-placing probability satisfies the condition

$$\frac{\mu-\eta}{\Lambda(2-\alpha)} \le q_J \le \frac{\mu-\eta}{\Lambda}. \quad \text{(A.3)}$$

The expected utility of a joining customer is $U_J^s(q_J,q_o) = q_o U_o^s + (1-q_o)U_s^s = V - P - \frac{C_h}{1-\alpha}$, which is independent of q_J.

(1) When $\frac{V-P}{C_h} < \frac{1}{1-\alpha}$ or equivalently, $\frac{(V-P)\alpha}{C} < \frac{1}{\eta}$, we have $U_J^s(q_J,q_o) < 0$ for all $q_J \in [0,1]$, so the best response of a customer is balking which contradicts to condition (A.3). Thus, the equilibrium does not exist.

(2) When $\frac{V-P}{C_h} \ge \frac{1}{1-\alpha}$ or equivalently, $\frac{(V-P)\alpha}{C} \ge \frac{1}{\eta}$, we have $U_J^s(q_J,q_o) \ge 0$ for all $q_J \in [0,1]$, so that $q_J = 1$ is an equilibrium if $\frac{\mu-\eta}{\Lambda(2-\alpha)} \le 1 \le \frac{\mu-\eta}{\Lambda}$ or equivalently, $\frac{1}{\eta} \le \frac{1}{\mu-(2-\alpha)\Lambda}$ and $\frac{1}{\mu-\Lambda} \le \frac{1}{\eta}$, which corresponds to the fourth case in Eq. (7).

(a.iii) If $q_o^s = 1$, then customers' order-placing probability satisfies the condition

$$q_J < \frac{\mu-\eta}{(2-\alpha)\Lambda}. \quad \text{(A.4)}$$

The expected utility of a joining customer is $U_J^s(q_J,q_o) = U_o^s(q_J,q_o) = \bar{V} - P - \frac{(2-\alpha)C}{\mu-\Lambda q_J(2-\alpha)}$. Because $U_J^s(q_J,q_o)$ is decreasing in q_J, we have the following results:

(1) When $\frac{\bar{V}-P}{(2-\alpha)C} \le \frac{1}{\mu}$, we have $U_J^s(q_J,q_o) \le 0$ for all $q_J \in [0,1]$, so the best response is to balk, which gives $q_J = 0 < \frac{\mu-\eta}{(2-\alpha)\Lambda}$, or equivalently, $\frac{1}{\mu} < \frac{1}{\eta}$.

(2) When $\frac{1}{\mu} < \frac{\bar{V}-P}{(2-\alpha)C} < \frac{1}{\mu-(2-\alpha)\Lambda}$, there exists a unique solution $\hat{q}_J^s$ to equation $U_J^s(q_J,q_o) = 0$ and it is an equilibrium if $\hat{q}_J^s = \frac{\mu - \frac{(2-\alpha)C}{\bar{V}-P}}{(2-\alpha)\Lambda} < \frac{\mu-\eta}{(2-\alpha)\Lambda}$, or equivalently, $\frac{\bar{V}-P}{(2-\alpha)C} < \frac{1}{\eta}$.

(3) When $\frac{\bar{V}-P}{(2-\alpha)C} \ge \frac{1}{\mu-(2-\alpha)\Lambda}$, we have $U_J^s(q_J,q_o) \ge 0$ for all $q_J \in [0,1]$, the best response is to place an order, which gives $q_J = 1 < \frac{\mu-\eta}{(2-\alpha)\Lambda}$, or equivalently, $\frac{1}{\mu-(2-\alpha)\Lambda} < \frac{1}{\eta}$.

We then summarize all the equilibria with $q_J^s = 0$ to the **0** case.

In Case (b) (i.e., the market size $\mu/(2-\alpha) \le \Lambda < \mu$), there are three sub-cases:

(b.i) If $q_o^s = 0$, similar to **Case** (a), the order-placing probability satisfies Condition (A.2). A joining customer's expected utility is $U_J^s(q_J, q_o) = U_s^s(q_J, q_o) = V - P - C_h - \frac{C}{\mu - \Lambda q_J}$, where $U_J^s(q_J, q_o)$ is decreasing in $q_J \in [0,1]$. Again, there are three cases specified by the value of q_J. Because the discussions are very similar to Case (a.i), we omit them here.

(b.ii) If $q_o^s \in (0,1)$, then customers' order-placing probability satisfies Condition (A.3). The expected utility of a joining customer is $U_J^s(q_J, q_o) = q_o U_o^s + (1-q_o) U_s^s = V - P - \frac{C_h}{1-\alpha}$, which is independent of q_J.

(1) When $\frac{V-P}{C_h} < \frac{1}{1-\alpha}$, we have $U_J^s(q_J, q_o) < 0$ for all $q_J \in [0,1]$, so the equilibrium joining probability is $q_J = 0$, contradicting to Condition (A.3). Thus, the equilibrium does not exist.

(2) When $\frac{V-P}{C_h} \ge \frac{1}{1-\alpha} \Leftrightarrow \frac{(V-P)\alpha}{C} \ge \frac{1}{\eta}$, we have $U_J^s(q_J, q_o) \ge 0$ for all $q_J \in [0,1]$, so $q_J = 1$ is an equilibrium if $\frac{\mu-\eta}{\Lambda(2-\alpha)} \le 1 \le \frac{\mu-\eta}{\Lambda}$, or equivalently, $\frac{1}{\eta} \le \frac{1}{\mu-(2-\alpha)\Lambda}$ and $\frac{1}{\mu-\Lambda} \le \frac{1}{\eta}$, while the first inequality contradicts the condition in Case (b). Thus, the equilibrium does not exist.

(b.iii) If $q_o^s = 1$, then customers' order-placing probability satisfies Condition (A.4). A joining customer's expected utility is $U_J^s(q_J, q_o) = U_o^s(q_J, q_o) = \bar{V} - P - \frac{(2-\alpha)C}{\mu - \Lambda q_J(2-\alpha)}$. Because $U_J^s(q_J, q_o)$ is decreasing in q_J, we have the following results:

(1) When $\frac{\bar{V}-P}{(2-\alpha)C} \le \frac{1}{\mu}$, we have $U_J^s(q_J, q_o) \le 0$ for $q_J \in [0,1]$, so the best response is balking, implying $q_J = 0 < \frac{\mu-\eta}{(2-\alpha)\Lambda}$, or equivalently, $\frac{1}{\mu} < \frac{1}{\eta}$.

(2) When $\frac{\bar{V}-P}{(2-\alpha)C} > \frac{1}{\mu}$, there exists a unique solution $\hat{q}_J^s$ to equation $U_J^s(q_J, q_o) = 0$ and $\hat{q}_J^s = \frac{\mu - \frac{(2-\alpha)C}{\bar{V}-P}}{(2-\alpha)\Lambda} < \frac{\mu-\eta}{(2-\alpha)\Lambda}$, or equivalently, $\frac{\bar{V}-P}{(2-\alpha)C} < \frac{1}{\eta}$.

In Case (c) (i.e., the market size $\Lambda \ge \mu$), there are three sub-cases:

(c.i) If $q_o^s = 0$, the order-placing probability satisfies the condition (A.2). A joining customer's expected utility is $U_J^s(q_J, q_o) = U_s^s(q_J, q_o) = V - P - C_h - \frac{C}{\mu - \Lambda q_J}$, where $U_J^s(q_J, q_o)$ is decreasing in $q_J \in [0,1]$.

(1) When $\frac{V-P-C_h}{C} \le \frac{1}{\mu}$, we have $U_J^s(q_J, q_o) \le 0$ for all $q_J \in [0,1]$, so the best response of all arriving customers is balking, indicating $q_J = 0 > \frac{\mu-\eta}{\Lambda}$, or equivalently, $\frac{1}{\mu} > \frac{1}{\eta}$.

(2) When $\frac{V-P-C_h}{C} > \frac{1}{\mu}$, there exists a unique solution $\tilde{q}_J^s$ to $U_J^s(q_J, q_o) = 0$, which is an equilibrium if $\tilde{q}_J^s = \frac{\mu - \frac{C}{V-P-C_h}}{\Lambda} > \frac{\mu-\eta}{\Lambda}$, or equivalently, $\frac{V-P-C_h}{C} > \frac{1}{\eta}$.

(c.ii) If $q_o^s \in (0,1)$, then customers' order-placing probability satisfies Condition (A.3). The expected utility of a joining customer is $U_J^s(q_J, q_o) = q_o U_o^s + (1-q_o) U_s^s = V - P - \frac{C_h}{1-\alpha}$, which is independent of q_J.

(1) When $\frac{V-P}{C_h} < \frac{1}{1-\alpha}$, then $U_J^s(q_J, q_o) < 0$ for all q_J, so the equilibrium joining probability is $q_J^s = 0$, contradicting Condition (A.3). Thus, the equilibrium does not exist.

(2) When $\frac{V-P}{C_h} \ge \frac{1}{1-\alpha}$, then $U_J^s(q_J, q_o) \ge 0$ for all q_J, so $q_J = 1$ is an equilibrium if $\frac{\mu-\eta}{\Lambda(2-\alpha)} \le 1 \le \frac{\mu-\eta}{\Lambda}$, or equivalently, $\frac{1}{\eta} \le \frac{1}{\mu-(2-\alpha)\Lambda}$ and $\frac{1}{\mu-\Lambda} \le \frac{1}{\eta}$, while the two inequalities contradict to the condition in Case (c). Hence, the equilibrium does not exist.

(c.iii) If $q_o^s = 1$, then customers' order-placing probability satisfies Condition (A.4). The expected utility of a joining customer is $U_J^s(q_J, q_o) = U_o^s(q_J, q_o) = \bar{V} - P - \frac{(2-\alpha)C}{\mu - \Lambda q_J(2-\alpha)}$. Hence, there are two cases specified by the value of q_J. Because the discussions are similar to (b.iii), we omit them here. □

Proof of Lemma 2. Using the customer utilities via the two service channels (12)–(13), we define the difference of these two functions as

$$\Delta^o(q_J, q_o) \equiv U_o^o(q_J, q_o) - U_s^o(q_J, q_o) = C_h - \frac{(1-\alpha)C}{\mu - \Lambda q_J[1 + (1-\alpha)q_o]}. \tag{A.5}$$

For a given q_J, the difference $\Delta^o(q_J, q_o)$ is decreasing in $q_o \in [0,1]$, and its maximum and minimum values are

$$\Delta^{\uparrow\uparrow}(q_J) \equiv \max_{q_o} \Delta^o(q_J, q_o) = \Delta^o(q_J, 0) = C_h - \frac{(1-\alpha)C}{\mu - \Lambda q_J},$$

$$\Delta^{\downarrow\downarrow}(q_J) \equiv \min_{q_o} \Delta^o(q_J, q_o) = \Delta^o(q_J, 1) = C_h - \frac{(1-\alpha)C}{\mu - \Lambda q_J(2-\alpha)}.$$

Define $\delta = \frac{(1-\alpha)C}{C_h}$, then we consider three cases:

(1) When $\frac{1}{\delta} < \frac{1}{\mu - \Lambda q_J}$, that is, $\Delta^{\uparrow\uparrow}(q_J) < 0$ and then $\Delta^o(q_J, q_o) < 0$ for all $q_o \in [0,1]$, thus the best response of a joining customer is to order onsite, which gives the equilibrium $q_o^o(q_J) = 0$;

(2) When $\frac{1}{\mu-\Lambda q_J} \le \frac{1}{\delta} \le \frac{1}{\mu - \Lambda q_J(2-\alpha)}$, there exists a unique solution q_o^* to $\Delta^o(q_J, q_o^*) = 0$, such that $\Delta^o(q_J, q_o) > 0$ for $q_o \in [0, q_o^*)$ and $\Delta^o(q_J, q_o) < 0$ for $q_o \in (q_o^*, 1]$. So the equilibrium order-selecting probability $q_o^o(q_J) = q_o^* = \frac{1}{1-\alpha}\left(\frac{\mu-\delta}{\Lambda q_J} - 1\right)$;

(3) When $\frac{1}{\delta} > \frac{1}{\mu - \Lambda q_J(2-\alpha)}$, that is, $\Delta^{\downarrow\downarrow}(q_J) > 0$ and then $\Delta^o(q_J, q_o) > 0$ for all $q_o \in [0,1]$, thus the best response of a joining customer is to order online, which gives the equilibrium $q_o^o(q_J) = 1$.

To prove the weakly decreasing property of q_o^o in q_J, we can pick two probabilities $q_J^1 < q_J^2$ and similar to the proof of Lemma 1, we find that $q_o^o(q_J^1) \ge q_o^o(q_J^2)$ always holds. □

Proof of Theorem 2. For **Case** (a), the market size $\Lambda < \mu/(2-\alpha)$, there are three sub-cases:

(a.i) If $q_o^o = 0$, recall the result in the first branch of Lemma 2, there must exists $\frac{1}{\delta} < \frac{1}{\mu - \Lambda q_J}$, which requires that the order-placing probability satisfies the condition

$$q_J > \frac{\mu - \delta}{\Lambda}. \tag{A.6}$$

Customers' expected utility of joining is $U_J^o(q_J, q_o) = U_s^o(q_J, q_o) = V - P - C_h - \frac{C}{\mu - \Lambda q_J}$, where $U_J^o(q_J, q_o)$ is decreasing in $q_J \in [0,1]$. Hence, three cases are discussed as follows:
(1) When $\frac{V-P-C_h}{C} \le \frac{1}{\mu}$, then $U_J^o(q_J, q_o) \le 0$ for all $q_J \in [0,1]$, so the best response of all arriving customers is to balk, with the equilibrium $q_J = 0 > \frac{\mu-\delta}{\Lambda}$, or equivalently, $\frac{1}{\mu} > \frac{1}{\delta}$. (2) When $\frac{1}{\mu} < \frac{V-P-C_h}{C} < \frac{1}{\mu-\Lambda}$, there exists a unique solution $\tilde{q}_J^o$ to $U_J^o(q_J, q_o) = 0$, which is an equilibrium if and only if $\tilde{q}_J^o = \frac{\mu - \frac{C}{V-P-C_h}}{\Lambda} > \frac{\mu-\delta}{\Lambda}$, or equivalently, $\frac{V-P-C_h}{C} > \frac{1}{\delta}$. (3) When $\frac{V-P-C_h}{C} \ge \frac{1}{\mu-\Lambda}$, then $U_J^o(q_J, q_o) \ge 0$ for all $q_J \in [0,1]$, so that joining is the best response for all customers and it is an equilibrium if and only if $1 > \frac{\mu-\delta}{\Lambda}$, or equivalently, $\frac{1}{\delta} < \frac{1}{\mu-\Lambda}$.

(a.ii) If $q_o^o \in (0,1)$, then customer's probability of joining satisfies the condition

$$\frac{\mu - \delta}{\Lambda(2-\alpha)} \le q_J \le \frac{\mu - \delta}{\Lambda}. \tag{A.7}$$

The expected utility of a joining customer is $U_J^o(q_J, q_o) = q_o U_o^o(q_J, q_o) + (1-q_o) U_s^o(q_J, q_o) = V - P - \frac{2-\alpha}{1-\alpha} C_h$, which is independent in q_J. There are two cases: (1) When $\frac{V-P}{C_h} < \frac{2-\alpha}{1-\alpha}$, then $U_J^o(q_J, q_o) < 0$ for all $q_J \in [0,1]$, so the best response of a customer is to balk, which contradicts to condition (A.7).

Thus, the equilibrium does not exist. (2) When $\frac{V-P}{C_h} \geq \frac{2-\alpha}{1-\alpha}$, or equivalently, $\frac{1}{\delta} \leq \frac{V-P}{(2-\alpha)C}$, then $U_J^o(q_J, q_o) \geq 0$ for all $q_J \in [0,1]$, so $q_J = 1$ is an equilibrium if and only if $\frac{\mu-\delta}{\Lambda(2-\alpha)} \leq 1 \leq \frac{\mu-\delta}{\Lambda}$, or equivalently, $\frac{1}{\delta} \leq \frac{1}{\mu-(2-\alpha)\Lambda}$ and $\frac{1}{\delta} \geq \frac{1}{\mu-\Lambda}$, which corresponds the fourth case in Eq. (15).

(a.iii) If $q_o^o = 1$, then the customer's probability of joining satisfies the condition

$$q_J < \frac{\mu-\delta}{(2-\alpha)\Lambda}. \tag{A.8}$$

The expected utility of joining is $U_J^o(q_J, q_o) = U_o^o(q_J, 1) = V - P - \frac{(2-\alpha)C}{\mu-\Lambda q_J(2-\alpha)}$. Using the fact that $U_J^o(q_J, q_o)$ is decreasing in q_J. We discuss the following results:
(1) When $\frac{V-P}{(2-\alpha)C} \leq \frac{1}{\mu}$, we have $U_J^o(q_J, q_o) \leq 0$ for $q_J \in [0,1]$, so the best response is to balk, which gives the equilibrium $q_J = 0 < \frac{\mu-\delta}{(2-\alpha)\Lambda}$, or equivalently, $\frac{1}{\mu} < \frac{1}{\delta}$. (2) When $\frac{1}{\mu} < \frac{V-P}{(2-\alpha)C} < \frac{1}{\mu-(2-\alpha)\Lambda}$, there exists a unique solution $\hat{q}_J^o$ to equation $U_J(q_J, q_o) = 0$ and $\hat{q}_J^o = \frac{\mu-\frac{(2-\alpha)C}{V-P}}{(2-\alpha)\Lambda} < \frac{\mu-\delta}{(2-\alpha)\Lambda}$, or equivalently, $\frac{V-P}{(2-\alpha)C} < \frac{1}{\delta}$. (3) When $\frac{V-P}{(2-\alpha)C} \geq \frac{1}{\mu-(2-\alpha)\Lambda}$, the best response is to place an order with $q_J^o = 1 < \frac{\mu-\delta}{(2-\alpha)\Lambda}$, or equivalently, $\frac{1}{\delta} > \frac{1}{\mu-(2-\alpha)\Lambda}$.

For **Case** (b), the market size $\mu/(2-\alpha) \leq \Lambda < \mu$, there are three sub-cases:

(b.i) If $q_o^o = 0$, similar to **Case** (a), the order-placing probability satisfies the condition (A.6). Customers' expected utility of joining is where $U_J^o(q_J, q_o)$ is decreasing in q_J. Hence, there are three cases about the value of q_J can be discussed, which are the same as **Case** (a.i).

(b.ii) If $q_o^o \in (0,1)$, then customers' order-placing probability satisfies the condition (A.7). The expected utility of a joining customer is $U_J^o(q_J, q_o) = q_o U_o^o + (1-q_o)U_s^o = V - P - \frac{(2-\alpha)C_h}{1-\alpha}$, which is independent of q_J. Hence, two sub-cases are discussed as follows: (1) When $\frac{V-P}{C_h} < \frac{2-\alpha}{1-\alpha}$, then $U_J^o(q_J, q_o) < 0$ for all $q_J \in [0,1]$, so the equilibrium joining probability is $q_J = 0$, which contradicts to condition (A.7). Thus, the equilibrium does not exist. (2) When $\frac{V-P}{C_h} \geq \frac{2-\alpha}{1-\alpha}$, or equivalently, $\frac{1}{\delta} \leq \frac{V-P}{(2-\alpha)C}$, then $U_J^o(q_J, q_o) \geq 0$ for all $q_J \in [0,1]$, so $q_J = 1$ is an equilibrium if and only if $\frac{\mu-\delta}{\Lambda(2-\alpha)} \leq 1 \leq \frac{\mu-\delta}{\Lambda}$, or equivalently, $\frac{1}{\delta} \leq \frac{1}{\mu-(2-\alpha)\Lambda}$ and $\frac{1}{\delta} \geq \frac{1}{\mu-\Lambda}$, while the first inequality contradicts to the condition in **Case** (b).

(b.iii) If $q_o^o = 1$, then customers' order-placing probability satisfies the condition (A.8). The expected utility of joining is $U_J^o(q_J, q_o) = U_o^o(q_J, 1) = V - P - \frac{(2-\alpha)C}{\mu-\Lambda q_J(2-\alpha)}$. Using the fact that $U_J^o(q_J, q_o)$ is decreasing in q_J, we have the following results:
(1) When $\frac{V-P}{(2-\alpha)C} \leq \frac{1}{\mu}$, we have $U_J^o(q_J, q_o) \leq 0$ for $q_J \in [0,1]$, so the best response is to balk, which gives $q_J = 0 < \frac{\mu-\delta}{(2-\alpha)\Lambda}$, or equivalently, $\frac{1}{\mu} < \frac{1}{\delta}$. (2) When $\frac{V-P}{(2-\alpha)C} > \frac{1}{\mu}$, there exists a unique solution $\hat{q}_J^o$ to equation $U_J(q_J, q_o) = 0$ and $\hat{q}_J^o = \frac{\mu-\frac{(2-\alpha)C}{V-P}}{(2-\alpha)\Lambda} < \frac{\mu-\delta}{(2-\alpha)\Lambda}$, or equivalently, $\frac{V-P}{(2-\alpha)C} < \frac{1}{\delta}$.

For **Case** (c), the market size $\Lambda \geq \mu$, there are three sub-cases:

(c.i) If $q_o^o = 0$, the order-placing probability satisfies the condition (A.6). Customers' expected utility of joining is where $U_J^o(q_J, q_o)$ is decreasing in $q_J \in [0,1]$. Hence, two sub-cases are discussed as follows: (1) When $\frac{V-P-C_h}{C} \leq \frac{1}{\mu}$, then $U_J^o(q_J, q_o) \leq 0$ for all $q_J \in [0,1]$, so the best response of all arriving customers is to balk, with $q_J = 0 > \frac{\mu-\delta}{\Lambda}$, or equivalently, $\frac{1}{\mu} > \frac{1}{\delta}$. (2) When $\frac{V-P-C_h}{C} > \frac{1}{\mu}$, there exists a unique solution $\tilde{q}_J^o$ to $U_J^o(q_J, q_o) = 0$, which is an equilibrium if and only if $\tilde{q}_J^o = \frac{\mu-\frac{C}{V-P-C_h}}{\Lambda} > \frac{\mu-\delta}{\Lambda}$, or equivalently, $\frac{V-P-C_h}{C} > \frac{1}{\delta}$.

(c.ii) If $q_o^o \in (0,1)$, then customers' order-placing probability satisfies the condition (A.7). The expected utility of a joining customer is $U_J^o(q_J, q_o) = V - P - \frac{(2-\alpha)C_h}{1-\alpha}$, which is independent of q_J. Hence, two cases are discussed as follows: (1) When $\frac{V-P}{C_h} < \frac{2-\alpha}{1-\alpha}$, then $U_J^o(q_J, q_o) < 0$ for all q_J, so the equilibrium joining probability is $q_J^o = 0$, which contradicts condition (A.7). Thus, the equilibrium does not exist. (2) When $\frac{V-P}{C_h} \geq \frac{2-\alpha}{1-\alpha}$, then $U_J^o(q_J, q_o) \geq 0$ for all q_J, so $q_J = 1$ is an equilibrium if and only if $\frac{\mu-\delta}{\Lambda(2-\alpha)} \leq 1 \leq \frac{\mu-\delta}{\Lambda}$, or equivalently, $\frac{1}{\delta} \leq \frac{1}{\mu-(2-\alpha)\Lambda}$ and $\frac{1}{\delta} \geq \frac{1}{\mu-\Lambda}$, while the two inequalities contradict to the condition in **Case** (c). The equilibrium does not exist.

(c.iii) If $q_o^o = 1$, then customers' order-placing probability satisfies the condition (A.8). The expected utility of joining is $U_J^o(q_J, q_o) = V - P - \frac{(2-\alpha)C}{\mu-\Lambda q_J(2-\alpha)}$. Hence, there are two cases about the value of q_J can be discussed, which are the same as (b.iii). □

Proof of Proposition 1. We use the equilibrium strategy in Theorem 1 to verify the impact of price P.

We first consider the *onsite exchange model*. The expected utility of a joining customer is

$$U_J^s(q_J^s, q_o^s) = q_o^s U_o^s + (1-q_o^s)U_s^s = V - P - (1-\alpha q_o^s)C_h - [1+(1-\alpha)q_o^s]CW(q_J^s, q_o^s),$$

which is decreasing in P. For any q_o^s, the equilibrium order-placing probability has the following piecewise structure, that is,

$$q_J^s = \begin{cases} 0, & \text{if} \quad P \geq \bar{V} - [1+(1-\alpha)q_o^s]CW(0, q_o^s), \\ q_J^*, & \text{if} \quad \bar{V} - [1+(1-\alpha)q_o^s]CW(1, q_o^s) < P < \bar{V} - [1+(1-\alpha)q_o^s]CW(0, q_o^s), \\ 1, & \text{if} \quad P \leq \bar{V} - [1+(1-\alpha)q_o^s]CW(1, q_o^s), \end{cases}$$

where $\bar{V} = V - (1-\alpha q_o^s)C_h$ and q_J^* uniquely solves by $U_J^s(q_J^s, q_o^s) = 0$. The joining probability q_J^s is a constant when P is sufficiently large (i.e., $P \geq \bar{V} - [1+(1-\alpha)q_o^s]CW(0, q_o^s)$) or sufficiently small (i.e., $P \leq \bar{V} - [1+(1-\alpha)q_o^s]CW(1, q_o^s)$). Otherwise, it is nonincreasing in P (with $q_J^s = q_J^*$). To see this, consider three cases: (1) if $q_o^s = 0$, q_J^*, as the solution to the equation $V - P - C_h - \frac{C}{\mu-\Lambda q_J} = 0$, decreases in P; (2) if $q_o^s = 1$, q^*J is the solution to the equation $\bar{V} - P - \frac{(2-\alpha)C}{\mu-\Lambda q_J(2-\alpha)} = 0$, which again is decreasing in P; (3) if $q_o^s \in (0,1)$, the joining utility is $U_J^s(q_J^s, q_o^s) = V - P - \frac{C_h}{1-\alpha}$ (which is decreasing in P), so that q_J^* is also decreasing in P. Similarly, it can be showed that the expected utility in the *online exchange model* of a joining customer is also decreasing in P.

In addition, invoking results in Lemmas 1 and 2, we have that q_o^s and q_o^o are nonincreasing in q_J^s and q_J^o respectively, which implies that the channel-selecting probability is nondecreasing in P. □

Proof of Proposition 2. We first work with the onsite-exchange model. First, we prove the monotonicity of the channel-selecting probability q_o^s with respect to the expectation-meeting probability α.

According to Lemma 1, the channel-selecting probability satisfies

$$q_o^s(q_J) = \begin{cases} 0, & \text{if} \quad \Lambda > \frac{\mu-\eta}{q_J}, \\ \frac{1}{1-\alpha}\left(\frac{\mu-\eta}{\Lambda q_J} - 1\right), & \text{if} \quad \frac{\mu-\eta}{(2-\alpha)q_J} \leq \Lambda \leq \frac{\mu-\eta}{q_J}, \\ 1, & \text{if} \quad \Lambda < \frac{\mu-\eta}{(2-\alpha)q_J}, \end{cases} \tag{A.9}$$

where $\eta = \frac{(1-\alpha)C}{\alpha C_h}$. To show that q_o^s is nondecreasing in α, we pick two values $\alpha_1 < \alpha_2$ and denote the corresponding channel-selecting probabilities as $q_o^{s,1}(\alpha_1)$ and $q_o^{s,2}(\alpha_2)$. Define $\eta^1 \equiv \frac{(1-\alpha_1)C}{\alpha_1 C_h}$, $\eta^2 \equiv \frac{(1-\alpha_2)C}{\alpha_2 C_h}$, $y^1 \equiv \frac{\mu-\eta^1}{q_J} < \frac{\mu-\eta^2}{q_J} \equiv y^2$, $x^1 \equiv \frac{\mu-\eta^1}{(2-\alpha_1)q_J} < \frac{\mu-\eta^2}{(2-\alpha_2)q_J} \equiv x^2$, and $z^1 \equiv \frac{1}{1-\alpha_1}\left(\frac{\mu-\eta^1}{\Lambda q_J} - 1\right) < \frac{1}{1-\alpha_2}\left(\frac{\mu-\eta^2}{\Lambda q_J} - 1\right) \equiv z^2$. We then discuss the following two cases:

(1) If $x^1 < x^2 \le y^1 < y^2$. Recall the equilibrium results in Eq. (A.9), when $\Lambda \le x^2$, then $q_o^{s,2}(\alpha_2) = 1 \ge q_o^{s,1}(\alpha_1)$; when $x^2 < \Lambda < y^1$, $q_o^{s,1}(\alpha_1) = z^1 < z^2 = q_o^{s,2}(\alpha_2)$; when $\Lambda \ge y^1$, $q_o^{s,2}(\alpha_2) \ge q_o^{s,1}(\alpha_1) = 0$.
(2) If $x^1 < y^1 < x^2 < y^2$. When $\Lambda \le x^2$, then $q_o^{s,2}(\alpha_2) = 1 \ge q_o^{s,1}(\alpha_1)$; when $\Lambda > x^2$, $q_o^{s,2}(\alpha_2) \ge q_o^{s,1}(\alpha_1) = 0$.

In summary, we have $q_o^{s,1}(\alpha_1) \le q_o^{s,2}(\alpha_2)$ for $\alpha_1 < \alpha_2$.

We now show that q_J^s is not monotonic in α. We first consider the small-market case. According to Theorem 1, we have that

$$(q_J^s, q_o^s) = \begin{cases} (\widetilde{q}_J^s, 0), & \text{if} \quad \max\left\{\frac{1}{\mu}, \frac{1}{\eta}\right\} \le \frac{V-P-C_h}{C} \le \frac{1}{\mu-\Lambda} \quad \text{(Condition 1)}, \\ (\widehat{q}_J^s, 1), & \text{if} \quad \frac{1}{\mu} \le \frac{\bar{V}-P}{(2-\alpha)C} \le \min\left\{\frac{1}{\mu-(2-\alpha)\Lambda}, \frac{1}{\eta}\right\} \quad \text{(Condition 2)}. \end{cases} \tag{A.10}$$

where $1/\eta = \frac{\alpha C_h}{(1-\alpha)C}$ is increasing in α. In addition, there exists a unique cutoff threshold $\widetilde{\alpha}$ such that Condition 1 (Condition 2) holds when $\alpha < \widetilde{\alpha}$ ($\alpha \ge \widetilde{\alpha}$), where $\widetilde{\alpha}$ satisfies $\frac{1}{\eta} = \frac{V-P-C_h}{C}$. We next compare the values $\widetilde{q}_J^s$ and $\widehat{q}_J^s$. Note that $\widetilde{q}_J^s$ is independent of α. By taking the first-order derivative of $\widehat{q}_J^s$ with respect to α, we obtain that $\frac{\partial \widehat{q}_J^s}{\partial \alpha} = \frac{-f'(\alpha)(2-\alpha)\Lambda+(\mu-f(\alpha))\Lambda}{(2-\alpha)^2\Lambda^2} > 0$, where $f(\alpha) \equiv \frac{(2-\alpha)C}{V-P-(1-\alpha)C_h} > 0$ and $f'(\alpha) = \frac{-C(V-P-(1-\alpha)C_h)-(2-\alpha)CC_h}{(V-P-(1-\alpha)C_h)^2} < 0$, which indicates that customers' order-placing probability $\widehat{q}_J^s$ is increasing in α. In summary, we have that

$$q_J = \begin{cases} \widetilde{q}_J^s, & \text{if} \quad \alpha < \widetilde{\alpha}, \\ \widehat{q}_J^s(\alpha), & \text{if} \quad \alpha \ge \widetilde{\alpha}. \end{cases}$$

Because $\widetilde{q}_J^s$ is independent of α and $\widehat{q}_J^s(\widetilde{\alpha}) < \widetilde{q}_J^s$ and $\widehat{q}_J^s$ is increasing in α, we have $\widehat{q}_J^s(\alpha) > \widetilde{\alpha}$. In summary, the order-placing probability q_J^s drops at $\widetilde{\alpha}$ and then increases in α afterwards, concluding the non-monotonic relationship in α. (Also see Fig. 7 for a visual illustration.)

The proofs for the other two cases can be done by following an identical road map. When the market size is medium, we mimic (A.10) by writing

$$(q_J^s, q_o^s) = \begin{cases} (\widetilde{q}_J^s, 0), & \text{if} \quad \max\left\{\frac{1}{\mu}, \frac{1}{\eta}\right\} \le \frac{V-P-C_h}{C} \le \frac{1}{\mu-\Lambda} \quad \text{(Condition 3)}, \\ (\widehat{q}_J^s, 1), & \text{if} \quad \frac{1}{\mu} \le \frac{\bar{V}-P}{(2-\alpha)C} \le \frac{1}{\eta} \quad \text{(Condition 4)}, \end{cases}$$

in place of Conditions 1 and 2 in (A.10). And when the market size is large, we can write

$$(q_J^s, q_o^s) = \begin{cases} (\widetilde{q}_J^s, 0), & \text{if} \quad \max\left\{\frac{1}{\mu}, \frac{1}{\eta}\right\} \le \frac{V-P-C_h}{C} \quad \text{(Condition 5)}, \\ (\widehat{q}_J^s, 1), & \text{if} \quad \frac{1}{\mu} \le \frac{\bar{V}-P}{(2-\alpha)C} \le \frac{1}{\eta} \quad \text{(Condition 6)}. \end{cases}$$

The rest of the analysis is similar to the small-market case: We can show that there also exists a cut off value $\widetilde{\alpha}$ which solves the equation $\frac{1}{\eta} = \frac{V-P-C_h}{C}$. Customers order-placing probability drops at $\widetilde{\alpha}$ and then increases in α afterwards, exhibiting the non-monotonic structure in α.

The corresponding results for the online-exchange model can be established similarly. □

Proof of Proposition 3. In both exchange models, the utility difference between the online and onsite service channels are given in Eqs. (A.1) and (A.5), both of which are increasing in α and $\Delta^o(q_J, q_o) > \Delta^s(q_J, q_o)$ for any fixed (q_J, q_o). In addition, $\Delta^s(q_J, q_o)|_{\alpha=0} < 0$ and $\Delta^s(q_J, q_o)|_{\alpha=1} > 0$. Therefore, there exist two thresholds $\bar{\alpha}$ and $\underline{\alpha}$ ($0 < \underline{\alpha} < \bar{\alpha} < 1$), where $\bar{\alpha}$ solves the equation $\Delta^s(q_J, q_o) = 0$ and $\underline{\alpha}$ solves $\Delta^o(q_J, q_o) = 0$. We consider three cases:

(1) When $\alpha < \underline{\alpha}$, we have $\Delta^s(q_J, q_o) < \Delta^o(q_J, q_o) < 0$, which shows that ordering onsite achieves a higher utility so all joining customers choose the onsite service channel ($q_o^s = q_o^o = 0$). Hence, customers' utilities in these two models are identical, i.e., $U_J^s(q_J, 0) = V - P - CW(q_J, 0) = U_J^o(q_J, 0)$ (the specific exchange policy has no impact on system performance). To be specific, $\underline{\alpha}$ solves the equation $\Delta^o(q_J, 0) = 0$ and the corresponding order-placing probability q_J satisfies $U_J^o(q_J, 0) = U_o^o(q_J, 0) = U_s^o(q_J, 0) = V - P - C_h - (2-\alpha)CW(q_J, 0) = 0$, which yields $\underline{\alpha} = \frac{\bar{V}-P-2C_h}{V-P-C_h}$.
(2) When $\alpha > \bar{\alpha}$, we have $\Delta^o(q_J, q_o) > \Delta^s(q_J, q_o) > 0$, which implies that customers' utility of selecting the online channel is always higher than that of the onsite channel, so all joining customers choose the online channel ($q_o^s = q_o^o = 1$). Customers' utility in the two models satisfy: $U_J^s(q_J, 1) = V - P - (2-\alpha)CW(q_J, 1) - (1-\alpha)C_h \ge V - P - (2-\alpha)CW(q_J, 1) = U_J^o(q_J, 1)$. Therefore, the online exchange model attracts more customers and generate higher system throughput. Specifically, $\bar{\alpha}$ satisfies $\Delta^s(q_J, 1) = 0$ and the order-placing probability q_J solves the equation $U_J^s(q_J, 1) = U_o^s(q_J, 1) = U_s^s(q_J, 1) = V - P - C_h - CW(q_J, 1)$, and this solves $\bar{\alpha} = 1 - \frac{C_h}{V-P}$.
(3) When $\underline{\alpha} \le \alpha \le \bar{\alpha}$, the customers' channel-selecting probability highly depends on the market size Λ.
   (3a) We first consider the case that the market size Λ is sufficiently small, which ensures that $U_J^s(1, q_o) \ge 0$ and $U_J^o(1, q_o) \ge 0$. While we have that $U_J^s(1, q_o) \ge 0 \Leftrightarrow V - P + (\alpha q_o - 1)C_h \ge [1 - q_o + (2-\alpha)q_o]CW(1, q_o)$ and $U_s^s(1, q_o) \ge 0 \Leftrightarrow V - P + (q_o - 1)C_h \ge [1 - q_o + (2-\alpha)q_o]CW(1, q_o)$. Hence, the maximum market size satisfying the conditions $U_J^s(1, q_o) \ge 0$ and $U_J^o(1, q_o) \ge 0$ can be solved from the equation $U_J^s(1, q_o) = 0$, which yields $\Lambda(q_o) = \frac{\mu - \frac{(q_o(2-\alpha)+1-q_o)C}{V-P-(1-\alpha q_o)C_h}}{1+(1-\alpha)q_o}$. In addition, because $U_J^s(1, q_o)$ is decreasing in q_o, the maximum value of Λ which satisfies the condition ($U_J^s(1, q_o) = 0$) is $\underline{\Lambda} = \Lambda(1) = \frac{\mu}{2-\alpha} - \frac{C}{V-P-(1-\alpha)C_h}$. When the market size $\Lambda \le \underline{\Lambda}$, all customers place orders in both exchange models ($q_J^s = q_J^o = 1$), resulting in identical system throughput.
   (3b) When the market size is large with $\Lambda > \underline{\Lambda}$, according to Proposition 1, the channel-selecting probability q_o^o and q_o^s are nondecreasing in price P. Therefore, there exists a threshold $\bar{P}$ for the price such that, when $P \ge \bar{P}$, $\Delta^o(q_J, q_o) > \Delta^s(q_J, q_o) > 0$, and $\bar{P}$ solves by $U_J^s(q_J, 1) = 0$ and $\Delta^s(q_J, 1) = 0$, and this gives $\bar{P} = V - C_h - \frac{\alpha C_h}{1-\alpha}$. Then all joining customers select the online channel ($q_o^s = q_o^o = 1$). This indicates that the online-exchange model outperforms the onsite-exchange model (analogous to case (1)). When $P < \bar{P}$, the utility difference satisfies $\Delta^o(q_J, q_o) > 0 > \Delta^s(q_J, q_o)$. Therefore, customers in the onsite-exchange model begin to switch to the onsite channel ($q_o^s < 1$), which enables the online channel selecting probability in the onsite-exchange model is lower than it in the online-exchange model. Hence, the corresponding joining probability q_J (and the system throughput) in the onsite-exchange model is higher, since we have proved the weakly decreasing property between q_J and q_o. In addition, the corresponding joining probability q_J solves $\Delta^s(q_J, q_o) = 0$. □

Proofs of Theorems 3 and 4. We first consider the onsite-exchange model. According to customers' equilibrium strategy in Theorem 1, we have $\widetilde{q}_J^s = 0$ if $\frac{1}{\mu} = \frac{V-P-C_h}{\mu}$, $\widetilde{q}_J^s = 1$ if $\frac{1}{\mu-\Lambda} = \frac{V-P-C_h}{C}$; and $\widehat{q}_J^s = 0$ if $\frac{1}{\mu} = \frac{\bar{V}-P}{(2-\alpha)C}$, $\widehat{q}_J^s = 1$ if $\frac{1}{\mu-(2-\alpha)\Lambda} = \frac{\bar{V}-P}{(2-\alpha)C}$. To facilitate the derivation of the optimal pricing strategy, we regroup the results in Theorem 1 we can present the equilibrium strategies in the following simplified format:

$$(q_J^s, q_o^s) = \begin{cases} (\widetilde{q}_J^s, 0), & \text{if} \quad \max\left\{\frac{1}{\mu}, \frac{1}{\eta}\right\} \le \frac{V-P-C_h}{C} \le \frac{1}{\mu-\Lambda}, \\ (1, \widehat{q}_o^s), & \text{if} \quad \frac{1}{\mu-\Lambda} \le \frac{1}{\eta} \le \min\left\{\frac{1}{\mu-(2-\alpha)\Lambda}, \frac{(V-P)\alpha}{C}\right\}, \\ (\widehat{q}_J^s, 1), & \text{if} \quad \frac{1}{\mu} \le \frac{\bar{V}-P}{(2-\alpha)C} \le \min\left\{\frac{1}{\mu-(2-\alpha)\Lambda}, \frac{1}{\eta}\right\}. \end{cases} \tag{A.11}$$

We remark that (A.11) is equivalent to the 6-branch structure as narrated in Theorem 1. We then discuss the following cases:

Case (1) $(q_J^s, q_o^s) = (\widetilde{q}_J^s, 0)$, the revenue-maximization problem is

$$\max_{P\geq 0} \Pi_s = \Lambda q_J^s P = \left(\mu - \frac{C}{V-P-C_h}\right) P \tag{A.12}$$
$$s.t. \quad \max\left\{\frac{1}{\mu}, \frac{1}{\eta}\right\} \leq \frac{V-P-C_h}{C} \leq \frac{1}{\mu-\Lambda}.$$

Taking the first- and second-order derivatives of Π_s with respect to P yields

$$\frac{\partial \Pi_s}{\partial P} = \mu - \frac{C(V-C_h)}{(V-P-C_h)^2} \quad \text{and} \quad \frac{\partial^2 \Pi_s}{\partial P^2} = -\frac{2C(V-C_h)}{(V-P-C_h)^3} < 0.$$

When $\mu - \frac{C}{V-C_h} < 0$, we have $\frac{\partial \Pi_s}{\partial P} < 0$; otherwise, setting $\frac{\partial \Pi_s}{\partial P} = 0$ gives the maximum point $\widetilde{P}_s = V - C_h - \sqrt{\frac{C(V-C_h)}{\mu}}$. The corresponding constraint of problem (A.12) is equivalent to $V - C_h - \frac{C}{\mu-\Lambda} \leq P \leq \min\left\{V - C_h - \frac{C}{\mu}, V - C_h - \frac{C}{\eta}\right\}$, so the optimal price P_s^* can be derived as follows:

(1a) If $\mu - \frac{C}{V-C_h} < 0$, the function Π_s decreases in P and $P_s^* = V - C_h - \frac{C}{\mu-\Lambda}$;

(1b) If $\mu - \frac{C}{V-C_h} \geq 0$ and $V - C_h - \frac{C}{\mu} \geq V - C_h - \frac{C}{\eta}$, there are three cases: (1b-i) when $\widetilde{P}_s \leq V - C_h - \frac{C}{\mu-\Lambda}$, the function Π_s decreases in P, so $P_s^* = V - C_h - \frac{C}{\mu-\Lambda}$; (1b-ii) when $V - C_h - \frac{C}{\mu-\Lambda} < \widetilde{P}_s < V - C_h - \frac{C}{\eta}$, the function Π_s is concave in P, then $P_s^* = \widetilde{P}_s$; (1b-iii) when $\widetilde{P}_s \geq V - C_h - \frac{C}{\eta}$, the function Π_s increases in P, then $P_s^* = V - C_h - \frac{C}{\eta}$.

(1c) If $\mu - \frac{C}{V-C_h} \geq 0$ and $V - C_h - \frac{C}{\mu} < V - C_h - \frac{C}{\eta}$, there are three cases: (1c-i) when $\widetilde{P}_s \leq V - C_h - \frac{C}{\mu-\Lambda}$, the function Π_s decreases in P, so $P_s^* = V - C_h - \frac{C}{\mu-\Lambda}$; (1c-ii) when $V - C_h - \frac{C}{\mu-\Lambda} < \widetilde{P}_s < V - C_h - \frac{C}{\mu}$, the function Π_s is concave in P, then $P_s^* = \widetilde{P}_s$; (1c-iii) when $\widetilde{P}_s \geq V - C_h - \frac{C}{\mu}$, the function Π_s increases in P then $P_s^* = V - C_h - \frac{C}{\mu}$.

Case (2) $(q_J^s, q_o^s) = (1, \widehat{q}_o^s)$, the revenue maximization problem is

$$\max_{P\geq 0} \Pi_s = \Lambda q_J^s P = \Lambda P, \tag{A.13}$$
$$s.t. \quad \frac{1}{\mu-\Lambda} \leq \frac{1}{\eta} \leq \min\left\{\frac{1}{\mu-(2-\alpha)\Lambda}, \frac{(V-P)\alpha}{C}\right\},$$

The revenue function Π_s is increasing in P and the constraint of problem (A.13) reduces to $P \leq V - \frac{C}{\eta\alpha}$, which concludes that $P_s^* = V - \frac{C}{\eta\alpha}$.

Case (3) $(q_J^s, q_o^s) = (\widehat{q}_J^s, 1)$, the revenue-maximization problem is

$$\max_{P\geq 0} \Pi_s = \Lambda q_J^s P = \frac{\mu - \frac{(2-\alpha)C}{\bar{V}-P}}{(2-\alpha)} P, \tag{A.14}$$
$$s.t. \quad \frac{1}{\mu} \leq \frac{\bar{V}-P}{(2-\alpha)C} \leq \min\left\{\frac{1}{\mu-(2-\alpha)\Lambda}, \frac{1}{\eta}\right\}.$$

Taking the first- and second-order derivatives of Π_s with respect to P yields

$$\frac{\partial \Pi_s}{\partial P} = \frac{1}{2-\alpha}\left(\mu - \frac{(2-\alpha)C\bar{V}}{(\bar{V}-P)^2}\right) \quad \text{and} \quad \frac{\partial^2 \Pi_s}{\partial P^2} = -\frac{2C\bar{V}}{(\bar{V}-P)^3} < 0.$$

When $\mu - \frac{(2-\alpha)C}{V-(1-\alpha)C_h} < 0$, we have $\frac{\partial \Pi_s}{\partial P} < 0$ and the revenue decreases in P; otherwise, setting $\frac{\partial \Pi_s}{\partial P} = 0$ gives the maximum point $\widehat{P}_s = \bar{V} - \sqrt{\frac{(2-\alpha)C\bar{V}}{\mu}}$, and the constraint of problem (A.14) becomes $\max\left\{\bar{V} - \frac{(2-\alpha)C}{\eta}, \bar{V} - \frac{(2-\alpha)C}{\mu-(2-\alpha)\Lambda}\right\} \leq P \leq \bar{V} - \frac{(2-\alpha)C}{\mu}$. The optimal price P_s^* are derived as follows:

(3a) If $\mu - \frac{(2-\alpha)C}{V-(1-\alpha)C_h} < 0$, and the function Π_s decreases in P, and P_s^* is the lower bound of the price. This implies that $P_s^* = \max\left\{\bar{V} - \frac{(2-\alpha)C}{\eta}, \bar{V} - \frac{(2-\alpha)C}{\mu-(2-\alpha)\Lambda}\right\}$.

(3b) If $\mu - \frac{(2-\alpha)C}{V-(1-\alpha)C_h} \geq 0$ and $\frac{(2-\alpha)C}{\eta} \leq \frac{(2-\alpha)C}{\mu-(2-\alpha)\Lambda}$, there exist three cases: (3b-i) when $\widehat{P}_s \leq \bar{V} - \frac{(2-\alpha)C}{\eta}$, the function Π_s is decreasing in P, $P_s^* = \bar{V} - \frac{(2-\alpha)C}{\eta}$; (3b-ii) when $\bar{V} - \frac{(2-\alpha)C}{\eta} < \widehat{P}_s < \bar{V} - \frac{(2-\alpha)C}{\mu}$, the function Π_s is concave in P, thus $P_s^* = \widehat{P}_s$; (3b-iii) when $\widehat{P}_s \geq \bar{V} - \frac{(2-\alpha)C}{\mu}$, the function Π_s is increasing in P, thus we have $P_s^* = \bar{V} - \frac{(2-\alpha)C}{\mu}$;

(3c) If $\mu - \frac{(2-\alpha)C}{V-(1-\alpha)C_h} \geq 0$ and $\frac{(2-\alpha)C}{\eta} > \frac{(2-\alpha)C}{\mu-(2-\alpha)\Lambda}$, there exist three cases: (3c-i) when $\widehat{P}_s \leq \bar{V} - \frac{(2-\alpha)C}{\mu-(2-\alpha)\Lambda}$, the function Π_s is decreasing in P, then $P_s^* = \bar{V} - \frac{(2-\alpha)C}{\mu-(2-\alpha)\Lambda}$; (3c-ii) when $\bar{V} - \frac{(2-\alpha)C}{\mu-(2-\alpha)\Lambda} < \widehat{P}_s < \bar{V} - \frac{(2-\alpha)C}{\mu}$, the function Π_s is concave in P, thus we have $P_s^* = \widehat{P}_s$; (3c-iii) when $\widehat{P}_s \geq \bar{V} - \frac{(2-\alpha)C}{\mu}$, the function Π_s is increasing in P, we have $P_s^* = \bar{V} - \frac{(2-\alpha)C}{\mu}$.

The proof of Theorem 4 is similar and is omitted here. □

Proof of Proposition 4. First, the system congestion increases in the market size Λ, leading to decreased customer utility ($U_J^s(q_J, q_o)$ and $U_J^o(q_J, q_o)$). Hence, the service provider needs to lower the service fee to attract consumers. Similar to the proof of Proposition 3, in the case of endogenous price, there also exist two thresholds for the expectation-meeting probability α of which the form are resemble those under the exogeneous price case (i.e., $\underline{\alpha}$ and $\bar{\alpha}$). Specifically, we now have $\underline{\alpha}_1 = \frac{V-\hat{P}_o^*-2C_h}{V-\hat{P}_o^*-C_h}$ and $\bar{\alpha}_1 = 1 - \frac{C_h}{V-\hat{P}_s^*}$, where $\hat{P}_o^*$ is the optimal price in the online-exchange model with $q_o^s = q_o^o = 1$ and $\hat{P}_s^*$ is the optimal price in the onsite-exchange model with $q_o^s = q_o^o = 0$. We next consider three cases:

(1) When $\alpha < \underline{\alpha}_1$, all joining customers order onsite (i.e., $q_o^s = q_o^o = 0$), leading to identical customer utility, behavior, system throughput and optimal revenue in the two exchange models. To be specific, $\hat{P}_o^*$ maximizes the revenue function $\Lambda q_J^o P$, where q_J^o satisfies $V - P - C_h - (2-\alpha)CW(q_J^o, 0) = 0$ (recall the proof in Proposition 3), and this gives $\hat{P}_o^* = V - C_h - \sqrt{\frac{C(V-C_h)}{\mu}}$ and then $\underline{\alpha}_1$ solves the equation $C_h^2\mu = (2-\alpha)(1-\alpha)^2C(V-C_h)$.

(2) When $\alpha > \bar{\alpha}_1$, all joining customers order online (i.e., $q_o^s = q_o^o = 1$), and the online-exchange model outperforms onsite-exchange model by giving a higher system throughput. To be specific, $\hat{P}_s^*$ maximizes the revenue function $\Lambda q_J^s P$, where q_J^s satisfies $V - P - C_h - CW(q_J^s, 1) = 0$ (recall the proof in Proposition 3), and this gives $\hat{P}_s^* = V - C_h - \sqrt{\frac{C(V-C_h)}{\mu}}$ and then $\bar{\alpha}_1 = \frac{\sqrt{\mu C(V-C_h)}}{\mu C_h + \sqrt{\mu C(V-C_h)}}$

(3) When $\underline{\alpha}_1 \leq \alpha \leq \bar{\alpha}_1$, customers' channel-selection probability depends on the market size:

(3a) When the market size is sufficiently small, the utilities in the two models satisfy $U_J^s(q_J^s, q_o^s) > 0$ and $U_J^o(q_J^o, q_o^o) > 0$, and then all customers join the system with $q_J^s = q_J^o = 1$. Similar to the proof of Proposition 3, the maximum market size satisfying the conditions $U_J^s(1, q_o) \geq 0$ and $U_J^o(1, q_o) \geq 0$ can be solved by the equation $U_J^s(1, q_o) = 0$, which yields $\Lambda(q_o) = \frac{\mu - \frac{(q_o(2-\alpha)+1-q_o)C}{V-\hat{P}_s^*-(1-\alpha q_o)C_h}}{1+(1-\alpha)q_o}$. And the lower bound of the market size is $\underline{\Lambda}' = \Lambda(1) = \frac{\mu}{2-\alpha} - \frac{\mu C}{\alpha\mu C_h + \sqrt{\mu C(V-C_h)}}$. Because in the online-exchange model, customers can harvest more surplus by sparing the inconvenience cost C_h, the service provider is able to achieve a higher revenue with a raised price.

(3b) As the market size increases ($\Lambda > \underline{\Lambda}'$), the utility difference satisfies $\Delta^o(q_J, q_o) > 0 > \Delta^s(q_J, q_o)$. Therefore, customers in the onsite-exchange model begin to switch to the on-site channel ($q_o^s < 1$), which enables $q_o^s < q_o^o$. Hence, the corresponding joining probability q_J (and the system throughput) in the onsite-exchange model is higher and then the onsite-exchange model begins to achieve a higher revenue. □

Proof of Theorem 5. Note that, the utility functions are decreasing in C_h. In addition, we have $U_o^s(p_s,p_o) < U_s^s(p_s,p_o)$ when $C_h = 0$. we consider two cases:

(1) When $\widetilde{C}_h \le \hat{C}_h$, we require the condition to hold:

$$W(p_s,p_o) = \frac{1}{\mu - \Lambda p_s - (2-\alpha)\Lambda p_o} \le \frac{\alpha(V-P)}{C}. \tag{A.15}$$

(1a) If $C_H < \bar{C}_h \Leftrightarrow W(p_s,p_o) > \frac{\alpha C_H}{(1-\alpha)C}$, all customers place onsite orders since $U_s^s(p_s,p_o) \ge U_o^s(p_s,p_o)$ and $U_s^s(p_s,p_o) \ge 0$, and $(p_s,p_o) = (1,0)$ is an equilibrium if and only if $\frac{\alpha C_H}{(1-\alpha)C} < W(1,0) = \frac{1}{\mu-\Lambda} \le \frac{\alpha(V-P)}{C}$.

(1b) If $\bar{C}_h \le C_H \le \hat{C}_h$, which is equivalent to conditions $W(p_s,p_o) \le \frac{\alpha C_H}{(1-\alpha)C}$ and $W(p_s,p_o) \le \frac{(V-P)-(1-\alpha)C_H}{(2-\alpha)C}$. Customers having $C_h \in [0,\bar{C}_h)$ will opt to the onsite service since $U_s^s(p_s,p_o) \ge U_o^s(p_s,p_o)$ and $U_s^s(p_s,p_o) \ge 0$, while customers with $C_h \in [\bar{C}_h, C_H]$ will place an online order because $U_s^s(p_s,p_o) < U_o^s(p_s,p_o)$ and $U_o^s(p_s,p_o) \ge 0$. Hence, $p_s = \bar{C}_h/C_H$ and $p_o = 1-p_s$, and p_s^1 satisfies $p_s^1 C_H = (1-\alpha)CW(p_s^1, 1-p_s^1)/\alpha$, and $(p_s,p_o) = (p_s^1, 1-p_s^1)$ is an equilibrium if and only if

$$W(p_s^1, 1-p_s^1) \le \min\left\{\frac{\alpha(V-P)}{C}, \frac{\alpha C_H}{(1-\alpha)C}, \frac{(V-P)-(1-\alpha)C_H}{(2-\alpha)C}\right\}.$$

(1c) If $C_H > \hat{C}_h \Leftrightarrow W(p_s,p_o) > \frac{(V-P)-(1-\alpha)C_H}{(2-\alpha)C}$, customers with $C_h \in [0,\bar{C}_h)$ will opt to order onsite since $U_s^s(p_s,p_o) \ge U_o^s(p_s,p_o)$ and $U_s^s(p_s,p_o) \ge 0$, and customers with $C_h \in [\bar{C}_h, \hat{C}_h)$ will place online orders since $U_s^s(p_s,p_o) < U_o^s(p_s,p_o)$ and $U_o^s(p_s,p_o) \ge 0$; otherwise, they balk since both $U_s^s(p_s,p_o)$ and $U_o^s(p_s,p_o)$ are negative. Hence, $p_s = \bar{C}_h/C_H$ and $p_o = (\hat{C}_h - \bar{C}_h)/C_H$, and $(p_s,p_o) = (p_s^2,p_s^2)$ is an equilibrium if an only if $\frac{(V-P)-(1-\alpha)C_H}{(2-\alpha)C} < W(p_s^2,p_s^2) \le \frac{\alpha(V-P)}{C}$.

(2) When $\widetilde{C}_h > \hat{C}_h$, we require the following condition to hold:

$$W(p_s,p_o) = \frac{1}{\mu - \Lambda p_s - (2-\alpha)\Lambda p_o} > \frac{\alpha(V-P)}{C}. \tag{A.16}$$

(2a) If $C_H \le \widetilde{C}_h \Leftrightarrow W(p_s,p_o) \le \frac{V-P-C_H}{C}$, all customers will choose onsite services since $U_s^s(p_s,p_o) \ge U_o^s(p_s,p_o)$ and $U_s^s(p_s,p_o) \ge 0$, and $(p_s,p_o) = (1,0)$ is an equilibrium if and only if $\frac{\alpha(V-P)}{C} < W(1,0) \le \frac{V-P-C_H}{C}$.

(2b) If $C_H > \widetilde{C}_h \Leftrightarrow W(p_s,p_o) > \frac{V-P-C_H}{C}$, customers with $C_h \in [0,\widetilde{C}_h)$ will place onsite orders because $U_s^s(p_s,p_o) \ge U_o^s(p_s,p_o)$ and $U_s^s(p_s,p_o) \ge 0$, and customers with $C_h \in [\widetilde{C}_h, C_H]$ balk since both $U_s^s(p_s,p_o)$ and $U_o^s(p_s,p_o)$ are negative. Hence, $p_s = \widetilde{C}_h/C_H$ and $p_o = 0$, and $(p_s,p_o) = (p_s^3, 0)$ is an equilibrium if and only if $W(p_s^3, 1-p_s^3) > \max\left\{\frac{\alpha(V-P)}{C}, \frac{V-P-C_H}{C}\right\}$.

The above cases correspond to the equilibrium in (24). □

Proof of Theorem 6. When $C_h = 0$, we have $U_o^o(p_s,p_o) < U_s^o(p_s,p_o)$, we consider two cases:

(1) When $U_o^o(p_s,p_o) \ge 0$, we need the following condition to hold:

$$V - P - (2-\alpha)CW(p_s,p_o) \ge 0 \Leftrightarrow W(p_s,p_o) \le \frac{V-P}{(2-\alpha)C}. \tag{A.17}$$

(1a) If $C_H < C_h' \Leftrightarrow W(p_s,p_o) > \frac{C_H}{(1-\alpha)C}$, all customers will choose onsite services because $U_s^o(p_s,p_o) \ge U_o^o(p_s,p_o)$ and $U_s^o(p_s,p_o) \ge 0$, so that $(p_s,p_o) = (1,0)$ is an equilibrium if and only if $\frac{C_H}{(1-\alpha)C} < W(1,0) \le \frac{V-P}{(2-\alpha)C}$.

(1b) If $C_H \ge C_h' \Leftrightarrow W(p_s,p_o) \le \frac{C_H}{(1-\alpha)C}$, customers with $C_h \in [0,C_h')$ will place onsite orders because $U_s^o(p_s,p_o) \ge U_o^o(p_s,p_o)$ and $U_s^o(p_s,p_o) \ge 0$, and customers with $C_h \in [C_h', C_H]$ will place online orders since $U_s^o(p_s,p_o) < U_o^o(p_s,p_o)$ and $U_o^o(p_s,p_o) \ge 0$. Hence, $p_s^1 = C_h'/C_H$ and $p_o^1 = 1-p_s^1$, with p_s^1 satisfying the equation $p_s^1 C_H = (1-\alpha)CW(p_1^s, 1-p_s^1)$, so that $(p_s,p_o) = (p_1^s, 1-p_s^1)$ is an equilibrium if and only if $W(p_s^1, 1-p_s^1) \le \min\left\{\frac{V-P}{(2-\alpha)C}, \frac{C_H}{(1-\alpha)C}\right\}$.

(2) When $U_o^o(p_s,p_o) < 0$, we require that

$$V-P-(2-\alpha)CW(q_J,q_o) < 0, \quad \text{or equivalently,} \quad W(q_J,q_o) > \frac{V-P}{(2-\alpha)C}. \tag{A.18}$$

(2a) If $C_H \le \underline{C}_h \Leftrightarrow W(p_s,p_o) \le \frac{V-P-C_H}{C}$, all customers will choose onsite services since $U_s^o(p_s,p_o) \ge U_o^o(p_s,p_o)$ and $U_s^o(p_s,p_o) \ge 0$, so that $(p_s,p_o) = (1,0)$ is an equilibrium if and only if $\frac{V-P}{(2-\alpha)C} < W(1,0) \le \frac{V-P-C_H}{C}$.

(2b) If $C_H > \underline{C}_h \Leftrightarrow W(p_s,p_o) > \frac{V-P-C_H}{C}$, customers with $C_h \in [0,\underline{C}_h)$ will place onsite orders and with $C_h \in [\underline{C}_h, C_H]$ balk. Hence, $p_s^3 = \underline{C}_h/C_H$ and $p_o^3 = 0$, so that $(p_s,p_o) = (p_s^3, 0)$ is an equilibrium if and only if $W(p_s^3, 0) > \max\left\{\frac{V-P}{(2-\alpha)C}, \frac{V-P-C_H}{C}\right\}$.

The above cases correspond to the equilibrium in (27). □

References

[1] Jopson B. US bricks-and-mortar retailers seek online success.. Financial Times 2013.
[2] Forrester R. Customer desires vs. retailer capabilities: minding the omni-channel commerce gap.. Technical report 2014.
[3] Zhang C, Zheng X. Customization strategies between online and offline retailers. Omega 2021;100:102230.
[4] Kuo CW, Xiong H, Chen YJ, Chang TK, Wu S, Chang YH. Vertical product line extension when online retailers serve as mom-and-pop stores' suppliers. Omega 2024;122:102944.
[5] Ranjan A, Misra M, Yadav DJ. Online shopping behavior during covid-19 pandemic: an indian perspective. Available at SSRN 3874348 2021.
[6] Levin A, Levin I, Heath C. Product category dependent consumer preferences for online and offline shopping features and their influence on multi-channel retail alliances.. Journal of Electronic Commerce Research 2003;4(3):85–93.
[7] Yang L, Li X, Xia Y, Aneja YP. Returns operations in omnichannel retailing with buy-online-and-return-to-store. Omega 2023;119:102874.
[8] BestBuy. Return and exchange policy. 2021, https://www.bestbuy.com/site/help-topics/return-exchange-policy/pcmcat260800050014.c?id=pcmcat260800050014.
[9] OldNavy. Return and exchange policy. 2021, https://oldnavy.gap.com/customerService/info.do?cid=82724.
[10] Ebay. Ebay, Return and Exchange Policies. 2021, https://www.thebalancesmb.com/refunds-and-exhcanges-for-ebay-purchases-1140366.
[11] Naor P. The regulation of queue size by levying tolls. Econometrica: Journal of the Econometric Society 1969;37(1):15–24.
[12] Edelson NM, Hilderbrand DK. Congestion tolls for poisson queuing processes. Econometrica: Journal of the Econometric Society 1975;43(1):81–92.
[13] Hassin R, Haviv M. To queue or not to queue: equilibrium behavior in queueing systems. 59, Springer Science & Business Media; 2003.
[14] Hassin R. Rational queueing. Boca Raton: CRC Press, Taylor and Francis Group; 2016.
[15] Economou A. The impact of information structure on strategic behavior in queueing systems. Chapter 4 in Anisimov, V. and Limnios, N. Queueing Theory 2: Advanced Trends. Wiley/ISTE.; 2021.
[16] Sun K, Liu Y, Liu X. Make waiting not to seem like waiting: capacity management of waiting-area entertainment.. Operations Research Letters 2022;50(6):745–52.
[17] Hassin R, Roet-Green R. The impact of inspection cost on equilibrium, revenue, and social welfare in a single-server queue. Operations Research 2017;65(3):804–20.
[18] Hassin R, Roet-Green R. On queue-length information when customers travel to a queue.. Manufacturing & Services Operations Management 2021;23(4):989–1004.
[19] Sun K, Liu Y, Yang L. Order ahead for pickup: promise or peril?. Available at SSRN 3673617 2022.
[20] Wang J, Sun K. Optimal pricing and capacity sizing for online service systems with free trials. OR Spectrum 2022;44(1):57–86.
[21] Choi TM, Guo S, Liu N, Shi X. Optimal pricing in on-demand-service-platform-operations with hired agents and risk-sensitive customers in the blockchain era. European Journal of Operational Research 2020;284(3):1031–42.

[22] Gao F, Su X. Omnichannel retail operations with buy-online-and-pick-up-in-store.. Management Science 2017;63(8):2478–92.
[23] Gao F, Su X. Omnichannel service operations with online and offline self-order technologies.. Management Science 2018;64(8):3595–608.
[24] Liu Y, Sun X, Hovey K. Scheduling to differentiate service in a multiclass service system. Operations Research 2022;70(1):527–44.
[25] Yang J, Huang J, Liu Y. Mind your own customers and ignore the others: asymptotic optimality of a local policy in multiclass queueing systems with customer feedback. IISE Transactions 2022;54(4):363–75.
[26] Roet-Green R, Yuan Y. Information visibility in omnichannel queues. Available at SSRN 3485810 2020.
[27] Chen M, Hu M, Wang J. Food delivery service and restaurant: friend or foe?. Management Science 2022;68(9):6539–51.
[28] He B, Gupta V, Mirchandani P. Online selling through O2O platform or on your own? Strategic implications for local brick-and-mortar stores. Omega 2021;103:102424.
[29] Zhou X, Liu H, Li J, Zhang K, Lev B. Channel strategies when digital platforms emerge: a systematic literature review. Omega 2023;120:102919.
[30] Baron O, Chen X, Li Y. Omnichannel services: the false premise and operational remedies. Management Science 2023;69(2):865–84.
[31] Guo P, Tang C, Wang Y, Zhao M. The impact of reimbursement policy on social welfare, revisit rate, and waiting time in a public healthcare system: fee-for-service versus bundled payment.. Manufacturing & Service Operations Management 2019;21(1):154–70.
[32] Aziz NNA, Wahid NA. Why consumers are hesitant to shop online: the major concerns towards online shopping. International Journal of Academic Research in Business and Social Sciences 2018;8(9):1175–85.
[33] Kumar A, Anjaly B. How to measure post-purchase customer experience in online retailing? A scale development study.. International Journal of Retail & Distribution Management 2017.
[34] Wang J, Wang Z, Chen Y. Equilibrium strategies and optimal pricing in an online retailing queueing system.. Naval Research Logistics 2021;68(5):556–76.
[35] Nageswaran L, Cho S, Scheller-Wolf A. Consumer return policies in omnichannel operations.. Management Science 2020;66(12):5558–75.
[36] Saleh. Starbucks prices products on value not cost. why?. 2020, https://thestrategystory.com/2020/09/16/starbucks-prices-products-on-value-not-cost-why.
[37] UberEats. Surge pricing is coming to ubereats in select cities. 2016, https://www.theverge.com/2016/10/27/13436334.