

Article

The Capacity Decision-Making of Omnichannel Catering Firms Based on Queueing System Considering Customer Reference Behavior

Wen-Tao Zhan ¹ , Xue-Ping Wang ¹, Ming-Hui Jiang ^{1,*}, Han Jiang ², Da Huo ¹ and Yun-Tao Liu ³¹ School of Management, Harbin Institute of Technology, Harbin 150001, China² Harbin Financial Service Bureau, Harbin 150001, China³ Nanyang Business School, Nanyang Technological University, Nanyang Avenue 639798, Singapore 308232, Singapore

* Correspondence: jiangminghui@hit.edu.cn

Abstract: Omnichannel catering firms need to pay attention to the consumer behavior of customers in both online and offline channels to make the corresponding capacity decisions. However, customers' consumption behavior is not only affected by base utility, but also by reference utility, which indicates that customers have reference dependence when making consumption decisions. In order to better help omnichannel catering firms to make capacity decisions more in line with customer consumption behavior in their actual operations, this paper adopts the queueing theory to construct an optimal capacity decision-making model for omnichannel catering firms. We also analyzed the impact of channel information availability on customer reference behavior and the subsequent impact on the firm's optimal decision-making. The findings were as follows: when the channel information is unavailable, customers in each channel take their expectation of waiting time as the reference point. With the improvement of the sensitivity of customers in different channels to the reference point, firms should improve their safety capacity to meet customers' needs. Moreover, at a higher reference point, the increase in customer sensitivity can make the firm obtain higher profits. When channel information is available, customers take the waiting time of customers in different channels as the reference point, and the increased sensitivity of online customers to the reference point will prompt firms to reduce safety capacity. Comparing the two scenarios, we also found that omnichannel catering firms could develop lower safety capacities and obtain higher profits more easily when channel information is available. This not only contributes to the development of omnichannel catering firms, but also expands the application scope of reference theory.

Keywords: reference behavior; queueing theory; omnichannel; capacity management

Citation: Zhan, W.-T.; Wang, X.-P.; Jiang, M.-H.; Jiang, H.; Huo, D.; Liu, Y.-T. The Capacity Decision-Making of Omnichannel Catering Firms Based on Queueing System Considering Customer Reference Behavior. *Systems* **2022**, *10*, 229. <https://doi.org/10.3390/systems10060229>

Academic Editor: Harish Garg

Received: 11 October 2022

Accepted: 17 November 2022

Published: 21 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The online channels have provided a huge impetus for the development of the catering industry. Especially with the rise of online platforms such as Meituan, Ubereats and Ele.me, as well as the advocacy of contactless services under the COVID-19, it seems to be the norm for customers to order online and wait for food delivery. In the traditional catering industry, firms are often offline single-channel, and customers pay attention to the waiting time, service environment and actual dining experience in physical stores. However, in omnichannel, merchants not only need to provide services to offline customers, but also need to provide services to online customers. This not only prompts firms to take into account the perception and demand of customers in multi-channels, but also puts forward requirements for the capacity decision-making of catering firms. In the process of capacity formulation, omnichannel catering firms must pay attention to the utility of customers and the change of demand. Some scholars often obtain the demand function through the net utility of customers [1,2], or directly provide the demand function of

customers to determine the demand [3–5]. Most of these scholars used the expected utility theory to describe customer utility, and they assumed that customers were completely rational. However, in the actual consumption process, customers tend to have referential behavior; that is, their perceived utility not only depends on the real utility of goods, but also depends on the relative amount of real utility and expected utility [6,7]. For example, when a customer enters a restaurant, he will compare the actual waiting time with the expected waiting time. When the actual waiting time is lower than the expected time, he will often get higher satisfaction from it. However, in contrast, once the actual waiting time is higher than the expectation, then this will cause customer dissatisfaction. This kind of relative quantity makes consumers have “loss” or “gain” psychology, which will affect the final utility of customers and cause a change in the customers’ purchase intention, so as to form the reference effect. However, once the firm ignores the customers’ reference behavior when developing the capacity of different stages, it will often cause an imbalance between the capacity and the demand, which will lead to loss of profits. In this regard, some scholars also started to study customer reference behavior [8–10]. Yang et al. [11] believe that the final utility of customers should also include reference utility, which will come from the comparison of utility and waiting time. After considering this reference utility, service providers can make better price and capacity decisions. Although these studies on customer reference behavior solve the problem of pricing decision in the operation of traditional firms, such customer reference behavior is still not considered in the operational process of omnichannel catering firms. This makes these catering firms often deviate from actual demand in the process of formulating capacities, which has an adverse impact on the development of such firms.

At the same time, we must also note that customers’ reference points will change with their perceptions [12]. Especially in omnichannel, when customers do not know the difference of customer utility between different channels, they cannot take the customer utility of other channels as the reference point. However, once customers are familiar with the utility of customers in other channels, they tend to take the utility of other customers as a reference point, which may lead to unfair avoidance psychology [13], resulting in the change of demand. Our study aims to solve how omnichannel catering firms should develop reasonable capacities to meet the needs of customers in different channels under the condition that customers have reference behavior. Therefore, the main problem to be solved in this paper is how will the reference behavior of customers in different channels affect their consumption? How should omnichannel catering firms make reasonable capacity decisions to meet the needs of different customers according to the reference behavior of customers? How will different reference points affect this capacity decision?

In order to solve these three problems, in this paper, we constructed a stylized decision model based on queuing theory and the reference behavior of customers, and analyzed the capacity decision of omnichannel firms through theoretical model analysis and numerical analysis. This could help omnichannel catering firms to make better reasonable capacity decisions in order to obtain optimal profit according to the actual perception and behavior of the customers in the system. We found that once the customers’ actual waiting time was shorter than their reference point, the improvement of the reference point sensitivity increased their shopping rate, which could further improve the profit of firms. In addition, when customers use their expectations as a reference point, the safety capacity of the firm increases with the sensitivity of the customer reference point. Moreover, when the customers’ reference point changes from expectation to waiting time of customers in different channels, the safety capacity established by the firm should be reduced instead.

Our innovation and contribution are three main points: Firstly, we introduce the reference behavior of customers into the capacity decision of omnichannel catering firms for analysis for the first time. This makes the decision of omnichannel merchants more in line with reality. Meanwhile, this analysis of the impact of customer reference behavior not only expands the application scope of the reference theory, but also enriches the omnichannel theory. Secondly, we study the impact of different customer reference points

on the optimal decision-making of firms according to the availability of information, and further explore the effect of the difference in reference points. Finally, we find that customer reference behavior always has an impact on the base capacities of catering firms, but it does not always have an impact on safety capacities. The relationship between optimal capacity and customer reference behavior presented in this paper not only fills the research gap of reference point effect in the field of capacity management, but also helps omnichannel catering firms to make more realistic decisions.

The remainder of this article is structured as follows. In Section 2, we review the relevant literature. In Section 3, we consider the optimal capacity decision of omnichannel catering firms when customers in different channels take their expected waiting time values as reference points when information is not available. In Section 4, we extend the model to the situation where information is available, and study the firms' optimal decisions when customers take the waiting time of customers in different channels as the reference point. In Section 5, we conduct numerical analysis based on the previous theoretical model. In Section 6, we draw some conclusions.

2. Literature Review

Our research is mainly related to two aspects of the literature, that is, customer reference behavior and omnichannel operation.

In terms of customer reference behavior, Kahneman and Tversky [14] first proposed prospect theory, where reference dependence is one of the important characteristics of individual choice behavior; that is, ordinary people's evaluation of a decision result is completed by calculating the change of the result relative to a reference point. People care about the difference between the final result and the reference point. On the basis of prospect theory and reference dependence, many scholars have conducted in-depth research into this kind of customer reference behavior. Among them, Blettner and Gollisch [15] studied the reference points used by managers in understanding performance feedback, which developed the application of reference points in organizational behavior. In terms of customer reference point selection, Popescu and Wu [8] took the exponential smoothing model as the price reference point and analyzed the pricing behavior of monopolies under dynamic conditions. Meanwhile, Wang et al. [16] improved the classic TOPSIS by adding a third intermediate reference point. They proposed three models of reference points, namely the model based on the best reference point, the model based on the worst reference point, and the model based on the classical best and worst reference point. Then, Qiu et al. [17] investigated the joint pricing and inventory decisions of loss-averse retailers with reference point effect under random demand, and found that loss-averse retailers with a smaller loss aversion coefficient benefited from higher reference effect strength. However, some operations management literature tended to take expected revenue as the reference point and considered both consumption utility and gain-loss utility [18,19]. Among them, Long and Nasiry [18] selected the reference point as the weighted average of the maximum and minimum possible income, and considered both consumption utility and gain-loss utility in the construction of the utility model. Yang et al. [11] made the optimal decision of the service operator based on the double reference point of price and waiting time, and found that for loss-averse customers, the firm could obtain higher profits in the duopoly market than in the monopoly market. However, we must note that the reference points selected by customers will change with the change of customers' cognition [12]. In particular, after knowing the utility and revenue of other customers, some customers tend to take the utility and revenue of other customers as their reference point to measure their own utility, because it is human nature to compare and weigh [20]. The above literature conducted relatively comprehensive and in-depth studies on the reference behavior of customers, but the customers in these works were usually one kind of customer without heterogeneity. Therefore, differently from these papers, this paper distinguished between the reference point effects generated by online and offline customers, and analyzed the impact of reference points on customer behavior in different channels, so as to make up

for the application of reference point effect on customer heterogeneity. At the same time, according to the availability of information, we also analyzed the impact of customers' choice of different reference points on firms' capacity decision-making, and expanded the variable reference point theory.

With regard to omnichannel study, Verhoef [21] gave an accurate definition of omnichannel marketing; that is, merchants pay attention to customer experience in the omnichannel, weakening the core position of physical stores, and organically integrating online and offline channels. Through cross-channel sales, customers can have an all-round consumption experience of shopping, entertainment and social interaction. Such an all-round consumption experience is important for customers, because they can make more convenient purchases on various apps [22]. Omnichannel provides customers with a variety of purchase methods, among which Gallino and Moreno [23] studied a new omnichannel practice from an empirical perspective, that is, buy online and pick up in store (BOPS), and they suggested that the implementation of BOPS may reduce online sales and increase offline traffic. Bell et al. [6,24] and Chopra [25] respectively studied the impact of "offline experience and online purchase" cooperation mode on consumer return behavior and firm operation efficiency by taking eyewear e-commerce Warby Parker and clothing e-commerce Bonobos as examples. With the development of the omnichannel, it also provides a new impetus for the catering industry. In omnichannel catering, delivery is also a concern of scholars. Yan et al. [26] believe that customers are sensitive to waiting time in the catering delivery system, and based on this, they built a food delivery network FooDNet, where the delivery time affected the experience of online customers. Therefore, Pei et al. [27] developed a method to evaluate online customer experience in the catering industry based on the BP neural network model, which solved the problem of measuring the actual customer experience. Zhao and Yang [28] studied the delivery service in the catering industry under the O2O mode based on the delivery time, and found that the self-delivery mode could help the delivery platform obtain higher profits. Chen et al. [29] investigated the long-term impact of food delivery service on the catering industry. They found that when the food delivery service is convenient enough, more customers using the food delivery service may harm the platform itself and society. This means that limiting the number of delivery workers provides a simple and effective way for platforms to boost their own profits while also benefiting society. However, there is still limited research into the capacity management of catering firms in omnichannel. Gao and Su [4] studied the impact of online and offline self-service technology on user demand, employment level and business interests in the ordering process, and analyzed the optimal capacity. They believe that although various technologies have been applied in the catering industry, firms still need to develop higher capacities to meet the needs of customers. Zhan et al. [5] analyzed the capacity development of merchants in the omnichannel catering industry in the ordering and production stages based on the influence of service environment, and they found that a better service environment could reduce customers' requirements for the production capacity of catering firms. Although this paper also focuses on the capacity management of catering firms, different from this kind of literature, we introduced the reference behavior of customers into omnichannel operations for the study. Then, we analyzed the impact of customer reference behavior in online and offline channels on the capacity management of omnichannel catering firms, and further studied how firms should adjust their capacities to meet customer needs under different customer reference points. This not only brings the capacity decisions of omnichannel catering firms more in line with the reality, but also bridges the research gap in the omnichannel field of customer reference behavior.

3. Channel Information Is Unavailable

3.1. Optimal Decision

Under different information availability conditions, customers' behaviors will be different, which will also affect the operation decisions of firms [30,31]. Therefore, we needed to investigate the degree of information acquisition by customers when analyzing

customers' behaviors. In this paper, we analyzed customer reference behavior and the operation decisions of omnichannel catering firms when channel information was available or unavailable. We need to state that the channel information we refer to is the waiting time of customers in other channels. In this section, we first study the case when channel information is unavailable.

We assume that an omnichannel catering firm provides food services to both online and offline customers, and it will provide a two-stage tandem system for offline customers. In stage 1, offline customers can place an order through a clerk at the front desk, and the order will be sent to the kitchen for food processing and preparation in stage 2. Offline customers need to wait during the ordering stage and food production stage. On the other hand, catering firms will provide online customers with a service process that includes online ordering, food production and delivery. For online customers, there is a need to wait during the food production stage and the delivery process, because they can order online without having to wait in line. However, under the current conditions of third-party delivery, omnichannel catering firms cannot decide the delivery time, so the delivery time t that online customers need to wait for will be an exogenous variable. In addition, the catering firm can get the revenue r from each customer. The rate μ_1 and μ_2 is the service capacity and production capacity in stage 1 and stage 2, respectively. Suppose that the cost of maintaining capacity μ_i at different stages $i = 1, 2$ is $c_i\mu_i$, where $c_i > 0$ is the cost of per unit capacity. This linear cost structure is widely used [4,5]. In order to ensure that it is profitable for the firm to serve each customer, it should be $r - c_1 - c_2 > 0$. The following are the notations of the model and their definition tables.

In this market, a single customer has no impact on the market. Meanwhile, the scale of customers in the whole market can be set as "1", that is, normalization is carried out, in which the proportion of offline customers is θ , while the proportion of online customers is $1 - \theta$. This approach to normalization of market size was applied in the study of Gao et al. [4] and Li et al. [32]. Customers have to wait in two stages before getting their food. $w_i(\mu_i, \Lambda)$ is the waiting time of customers at stage $i = 1, 2$ under given capacity μ_i and total demand Λ . Meanwhile, for each customer, there are expected waiting times \bar{W}_s and \bar{W}_o in the ordering stage and in the production stage. When channel information is unavailable, customers cannot know the waiting time of customers in other channels. Therefore, according to the reference theory proposed by Kahneman and Tversky [14], they take the expected waiting time as the reference point. Therefore, offline and online customers' reference points for time can be expressed as follows:

$$\text{reference point} = \begin{cases} \bar{W}_s & \text{offline customers} \\ \bar{W}_o & \text{online customers} \end{cases} \quad (1)$$

According to the above analysis, the actual waiting time of offline customers in the system is $W_s = w_1 + w_2$, while the actual waiting time of online customers is $W_o = w_2 + t$. Since customers have reference behavior, they will compare the actual waiting time with the expected waiting time, and their sensitivity to this reference difference can be expressed as δ_s and δ_o , respectively. When the actual waiting time is shorter than the reference point, customers obtain a higher utility from it and thus increase the shopping rate. In addition, the sensitivity of offline customers to waiting time is β . Since online customers can wait in the office or at home after ordering food online, their opportunity cost in the waiting process can be regarded as zero, with only anxiety cost [33]. The proportion of anxiety cost is ε , and $0 < \varepsilon < 1$, so the waiting cost per unit time of online customers is $\varepsilon\beta$. Since the shopping rate of customers decreases with the increase in waiting time, according to the linear equation of customer demand given by Gao and Su [4] and Zhan et al. [5], we can obtain the shopping rate of offline and online customers as follows (the subscript n indicates that the channel information is unavailable):

$$\lambda_{n,s} = [\alpha - \beta(w_1(\mu_1, \theta\lambda_s) + w_2(\mu_2, (\theta\lambda_s + (1 - \theta)\lambda_o))) - \delta_s(W_s - \bar{W}_s)]^+ \quad (2)$$

$$\lambda_{n,o} = [\alpha - \varepsilon\beta(w_2(\mu_2, (\theta\lambda_s + (1 - \theta)\lambda_o)) + t) - \delta_o(W_o - \bar{W}_o)]^+ \tag{3}$$

The above $[x]^+ = \max(x, 0)$, because the effective customer shopping rate must be positive in the actual operation [4,5]. The above α represents the base shopping rate of customers, which is assumed to be the same across different customers [4,5]. This base shopping rate is based on basic utility, and the basic utility obtained by customers when they purchase food depends on the quality and taste of food, which is the same for online and offline customers, so it can be considered that the base shopping rate of online and offline customers is the same. To simplify the analysis, we assume that the base shopping rate of customers in different channels is the same. The second term represents the impact on customer shopping rate caused by the change in waiting time. The third term describes the customers' reference behavior. When the actual waiting times W_s and W_o are smaller than \bar{W}_s and \bar{W}_o , respectively, customers obtain positive utility from the reference behavior, so as to improve their shopping rate.

On the other hand, the firm will choose the appropriate capacity to obtain the optimal profit according to the shopping rate of customers from different channels. The decision-making model is as follows:

$$\begin{aligned} & \max_{\theta\lambda_{n,s} < \mu_{n,1}, \theta\lambda_{n,s} + (1-\theta)\lambda_{n,o} < \mu_{n,2}} \{r(\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o}) - c_1\mu_{n,1} - c_2\mu_{n,2}\} \\ \text{s.t. } & \lambda_{n,s} = [\alpha - \beta(w_1(\mu_{n,1}, \theta\lambda_{n,s}) + w_2(\mu_{n,2}, (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o}))) - \delta_s(W_s - \bar{W}_s)]^+ \\ & \lambda_{n,o} = [\alpha - \varepsilon\beta(w_2(\mu_{n,2}, (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o})) + t) - \delta_o(W_o - \bar{W}_o)]^+ \end{aligned} \tag{4}$$

Although our model is built on the basis of Gao et al. [4] and Zhan et al. [5], different from their studies, our study introduces the reference behavior of customers in actual operation into the model of capacity decision-making, which makes the capacity decision of the catering firm more in line with the actual operation situation. By solving the above model, we can obtain the following proposition.

Proposition 1. *Under the premise of $r > c_1 + c_2$, the firm can choose the following capacities to provide services for customers in online and offline channels. (See Appendix A for proof)*

$$\mu_{n,1} = \theta\lambda_{n,s} + \sqrt{\frac{\theta(\beta + \delta_s)(r - c_1 - c_2)}{c_1}}, \tag{5}$$

$$\mu_{n,2} = \theta\lambda_{n,s} + (1 - \theta)\lambda_o + \sqrt{\frac{\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\varepsilon\beta + \delta_o)(r - c_2)}{c_2}}, \tag{6}$$

where

$$\lambda_{n,s} = \alpha - (\beta + \delta_s)\sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2)}} - (\beta + \delta_s)\sqrt{\frac{c_2}{\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\varepsilon\beta + \delta_o)(r - c_2)}} + \delta_s\bar{W}_s > 0, \tag{7}$$

$$\lambda_{n,o} = \alpha - (\varepsilon\beta + \delta_o)\sqrt{\frac{c_2}{\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\varepsilon\beta + \delta_o)(r - c_2)}} - (\varepsilon\beta + \delta_o)t + \delta_o\bar{W}_o > 0. \tag{8}$$

The above α needs to satisfy the range in the proof of Proposition 1 in the Appendix A. According to Proposition 1, the optimal capacity $\mu_{n,i}$, $i = 1, 2$ is composed of base capacity $\mu_{n,i}^b$ and safety capacity $\mu_{n,i}^{safe}$, which satisfies the rule of thumb of capacity planning as described by Bassamboo et al. [34]. Thus, we find the base capacity is $\mu_{n,1}^b = \theta\lambda_s$ and $\mu_{n,2}^b = \theta\lambda_s + (1 - \theta)\lambda_o$; the safety capacity is $\mu_{n,1}^{safe} = \sqrt{\frac{\theta(\beta + \delta_s)(r - c_1 - c_2)}{c_1}}$ and $\mu_{n,2}^{safe} = \sqrt{\frac{\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\varepsilon\beta + \delta_o)(r - c_2)}{c_2}}$. In the queuing system, the function of base capacity is mainly to meet the base customer demand, and its specific size will increase with the increase in customer demand. However, the main function of safety capacity is to hedge the demand changes of customers in the system, so as to ensure the safe operation of

the system, which is of great significance for firms. It was found that with the increase in c_i of unit capacity cost in different stages, the firm should reduce safety capacity to save more cost. However, this reduces the firm's ability to hedge against changes in market demand. For example, when the salary level of the kitchen staff is high, that is, the c_2 is relatively high, then the firm will not hire more kitchen staff, which will lead to the production capacity of the kitchen being insufficient, and unable to cope with the lunch or dinner peak demand of a large number of customers. From the above analysis, it can be seen that the capacities of the ordering stage and production stage will be related to each other through customer demand $\lambda_{n,s}$, unit capacity cost c_i , etc., rather than being independent.

3.2. Sensitivity Analysis

Lemma 1. *As offline customers become more sensitive to reference points, that is, δ_s increases, the omnichannel catering firm should set a higher level of ordering stage safety capacity and production stage safety capacity, i.e., $\frac{\partial \mu_{n,1}^{safe}}{\partial \delta_s} > 0$, $\frac{\partial \mu_{n,2}^{safe}}{\partial \delta_s} > 0$. However, the increase in online customers' sensitivity δ_o to the reference point does not affect the safety capacity in the ordering stage, but only increases the safety capability $\mu_{n,2}^{safe}$ in the production stage, i.e., $\frac{\partial \mu_{n,1}^{safe}}{\partial \delta_o} = 0$, $\frac{\partial \mu_{n,2}^{safe}}{\partial \delta_o} > 0$.*

From the above lemma, we can see that once the information between channels is unavailable, then customers do not know the waiting time of customers in other channels. In this case, customers take their expectation of waiting time as the reference point. The offline customers' sensitivity δ_s to the reference point gradually increases, which means that customers want the actual waiting time to be shorter. However, once the actual waiting time is longer than the expected waiting time, customers' consumption behavior will be inhibited. Therefore, the firm should enhance the safety capacity in the ordering stage and the production stage at the same time, so as to enhance the ability to hedge market demand changes and reduce the actual waiting time of customers. Moreover, the simultaneous improvement of capacity in the two stages is actually to ensure the coordination between the different stages.

However, for online customers, the improvement of their sensitivity δ_o does not cause the change of the ability in the ordering stage, because online customers use online ordering, but it does not occupy the service capacity of the firm in the ordering stage. However, the increase in δ_o still encourages firms to improve the safety capacities in the production stage to ensure that they can meet the needs of online customers. It can also be seen that when $\delta_s = 0$ and $\delta_o = 0$, the safety capacity of the firm satisfies $\mu_{n,1,\delta_s=0}^{safe} = \sqrt{\frac{\theta\beta(r-c_1-c_2)}{c_1}} < \mu_{n,1}^{safe}$ and $\mu_{n,2,\delta_o=0}^{safe} = \sqrt{\frac{\theta\beta(r-c_1-c_2)+(1-\theta)\varepsilon\beta(r-c_2)}{c_2}} < \mu_{n,2}^{safe}$. This indicates that if the omnichannel catering firm does not consider the reference behavior of customers, then the safety capacity is low and cannot meet the actual customers' needs.

Lemma 2. (1) *When the expected value \bar{W}_s is small, there is $\frac{\partial \lambda_{n,s}}{\partial \delta_s} < 0$ and $\frac{\partial \lambda_{n,s}}{\partial \delta_o} > 0$; When \bar{W}_s is large, $\frac{\partial \lambda_{n,s}}{\partial \delta_s} > 0$ and $\frac{\partial \lambda_{n,s}}{\partial \delta_o} > 0$ must be satisfied. (2) When \bar{W}_o is small, there is $\frac{\partial \lambda_{n,o}}{\partial \delta_o} < 0$, $\frac{\partial \lambda_{n,o}}{\partial \delta_s} > 0$; When \bar{W}_o is large, $\frac{\partial \lambda_{n,o}}{\partial \delta_o} > 0$ and $\frac{\partial \lambda_{n,o}}{\partial \delta_s} > 0$ must be satisfied. (3) It is obvious that the offline customers' shopping rate λ_s will increase with the rise of offline customers' reference points \bar{W}_s , i.e., $\frac{\partial \lambda_{n,s}}{\partial \bar{W}_s} > 0$. Similarly, there is $\frac{\partial \lambda_{n,o}}{\partial \bar{W}_o} > 0$ for the online customers' shopping rate. (See Appendix A for proof).*

Lemma 2 shows that higher customer sensitivity leads to higher purchase rates from other channels. The improvement of the customers' sensitivity to the reference point in any channel can improve the shopping rate of customers in another channel. In combination with Lemma 1, it can be seen that the increase in shopping rate is because the increase in δ_s or δ_o can prompt the firm to improve the safety capacity in the production stage, and the

improvement of safety capacity can comprehensively reduce the waiting time of customers in both online and offline channels, thus increasing their shopping rate.

When the expectation \bar{W}_s or \bar{W}_o is low, customers want a shorter waiting time, which puts forward higher requirements for the firm's capacity. Generally speaking, under the condition of fixed costs c_1 and c_2 , it is difficult for firms to improve their service and production capacity quickly, which means that customers cannot get positive utility from reference behavior. With the increase in sensitivities δ_s and δ_o , this disutility will be stronger, so that customers' shopping rate will decrease. In contrast, higher expected values of \bar{W}_s and \bar{W}_o will be more conducive for customers to increase their shopping rate, which is what is stated in Lemma 2 (3).

From Proposition 1, we can obtain the profits of catering firms as follows:

$$\pi_n = \theta(\alpha + \delta_s \bar{W}_s)(r - c_1 - c_2) + (1 - \theta)(\alpha - (\varepsilon\beta + \delta_o)t + \delta_o \bar{W}_o)(r - c_2) - 2\sqrt{c_1\theta(\beta + \delta_s)(r - c_1 - c_2)} - 2\sqrt{c_2(\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\varepsilon\beta + \delta_o)(r - c_2))}$$

Lemma 3. (1) When the actual waiting time satisfies $w_1 + w_2 > \bar{W}_s$, with the increase in offline customer sensitivity δ_s , profit will decrease, i.e., $\frac{\partial \pi_n}{\partial \delta_s} < 0$. In contrast, with the increase in δ_s , profit will increase, i.e., $\frac{\partial \pi_n}{\partial \delta_s} > 0$.

(2) For online customers, when the actual waiting time satisfies $w_2 + t > \bar{W}_o$, the increase in online customers' sensitivity δ_o will lead to the decrease in the firm's profit, i.e., $\frac{\partial \pi_n}{\partial \delta_o} < 0$. In contrast, the profit will increase with the increase in δ_o , i.e., $\frac{\partial \pi_n}{\partial \delta_o} > 0$. (See Appendix A for proof).

Lemma 3 shows that when the actual waiting time is shorter than the customer's expectation, the increase in customer sensitivity can increase the shopping rate of customers in online and offline channels, which can bring greater profits to the firm. Although the firm still needs to bear the cost increase brought in this case by the increase in safety capacity, the revenue increase brought by the increase in customer shopping rate can make up for the cost expenditure, so that the profits will increase. However, once the actual waiting time is longer than the customer's expectation, the customer experiences disutility from the reference behavior, and the customers' shopping rate will decrease. Especially, with the increase in sensitivity δ_o or δ_s , the disutility obtained by customers from the reference behavior will be strengthened, which is not conducive to profit.

As the actual waiting times $w_2 + t$ and $w_1 + w_2$ of online and offline customers are different, their expected waiting times \bar{W}_o and \bar{W}_s are also different. Thus, some cases will happen as follows:

- (1) $w_1 + w_2 > \bar{W}_s$ and $w_2 + t > \bar{W}_o$. Then the omnichannel catering firm's profit π_n increases with the increase in δ_s and δ_o .
- (2) $w_1 + w_2 > \bar{W}_s$ and $w_2 + t < \bar{W}_o$. The firm's profit π_n increases with the increase in δ_s , but decreases with the increase in δ_o .
- (3) $w_1 + w_2 < \bar{W}_s$, and $w_2 + t > \bar{W}_o$. The firm's profit π_n decreases with the increase in δ_s , but increases with the increase in δ_o .
- (4) $w_1 + w_2 < \bar{W}_s$ and $w_2 + t < \bar{W}_o$. The firm's profit π_n will not only decrease with the increase in δ_s , but also increase with the increase in δ_o .

The above cases ably demonstrate the impact of customer reference behavior on the profit of the omnichannel catering firm when information from other channels is unavailable. It can also clearly be learned that when customers of different channels have greater expectations of waiting time, that is, \bar{W}_s and \bar{W}_o are larger, the reference behavior of customers is conducive to profit. Once \bar{W}_s and \bar{W}_o are small, such customer reference behavior will harm profits, thus causing a bad impact on the development of the firm.

With the development of social networks and information networks, information between channels will gradually become known to customers in other channels. In this case, customers will change their reference points. In the next section, we conduct a detailed analysis of the impact of customer reference point changes on the capacity management and

profit of the omnichannel catering firm under the condition that information of different channels is available.

(When the actual waiting time is equal to the expected waiting time, the reference point effect of the customer disappears. Our paper will not continue to analyze this situation).

4. Channel Information Is Available

4.1. Optimal Decision

In the previous section, we studied the impact of online and offline customer reference behavior on the optimal capacity decision-making of the omnichannel catering firm when information from different channels is unavailable. However, with the development of modern communication means and the rise of various social platforms and networks, customers can effectively know the corresponding consumption information of customers in other channels, including waiting time information [3]. Once customers know the waiting time of customers in other channels, their reference points will often change [12]. In particular, after knowing the utility and revenue of other customers, some customers tend to take the utility and revenue of other customers as their reference point to measure their own utility, because it is human nature to compare and weigh [20]. For example, when offline customers know the waiting time W_o of online channel customers, they will compare their waiting time with that of online customers. In other words, the waiting time W_o of online customers becomes the reference point of offline customers. According to the reference theory, the shopping rate of each customer in the offline channel can be expressed as follows (the subscript a indicates that the channel information is available):

$$\lambda_{a,s} = [\alpha - \beta(w_1(\mu_1, \theta\lambda_s) + w_2(\mu_2, (\theta\lambda_s + (1 - \theta)\lambda_o))) - \delta_s(W_s - W_o)]^+ \quad (9)$$

According to the unfairness theory proposed by Fehr and Schmidt [13], that is, the perception of unfairness is based on the perception of unfairness generated when others get it but they fail to get it, once $W_o < W_s$, the reference point sensitivity δ_s , here, can actually be regarded as the unfairness perception coefficient of offline customers on waiting time. This also means that once the customer selects the waiting time of the customer in another channel as the reference point, the customer's reference behavior will be linked with fairness. Similarly, for online customers, their effective shopping rate can be expressed as follows:

$$\lambda_{a,o} = [\alpha - \varepsilon\beta(w_2(\mu_2, (\theta\lambda_s + (1 - \theta)\lambda_o)) + t) - \delta_o(W_o - W_s)]^+ \quad (10)$$

For the omnichannel catering firm, he will still formulate reasonable capacities according to the shopping rates of customers in the above different channels to obtain the optimal profits. The decision model can be expressed as follows:

$$\begin{aligned} & \max_{\theta\lambda_{a,s} < \mu_{a,1}, \theta\lambda_{a,s} + (1-\theta)\lambda_{a,o} < \mu_{a,2}} \{r(\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}) - c_1\mu_{a,1} - c_2\mu_{a,2}\} \\ \text{s.t. } & \lambda_{a,s} = [\alpha - \beta(w_1(\mu_{a,1}, \theta\lambda_{a,s}) + w_2(\mu_{a,2}, (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}))) - \delta_s(W_s - W_o)]^+ \\ & \lambda_{a,o} = [\alpha - \varepsilon\beta(w_2(\mu_{a,2}, (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})) + t) - \delta_o(W_o - W_s)]^+ \end{aligned} \quad (11)$$

Through solving the above equation, the optimal capacities of catering firms at different stages in the omnichannel can be expressed as follows:

Proposition 2. *In an omnichannel operation, when $r > c_1 + c_2$ is satisfied, the firm can choose the following capacity decisions to obtain the optimal profit (See Appendix A for proof):*

$$\mu_{a,1} = \theta\lambda_{a,s} + \sqrt{\frac{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}{c_1}}, \quad (12)$$

$$\mu_{a,2} = \theta\lambda_{a,s} + (1 - \theta)\lambda_o + \sqrt{\frac{\theta\beta(r - c_1 - c_2) + (1 - \theta)\varepsilon\beta(r - c_2)}{c_2}}, \tag{13}$$

where

$$\lambda_{a,s} = \alpha - (\beta + \delta_s)\sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}} - \beta\sqrt{\frac{c_2}{\theta\beta(r - c_1 - c_2) + (1 - \theta)\varepsilon\beta(r - c_2)}} + \delta_s t > 0, \tag{14}$$

$$\lambda_{a,o} = \alpha - \varepsilon\beta\sqrt{\frac{c_2}{\theta\beta(r - c_1 - c_2) + (1 - \theta)\varepsilon\beta(r - c_2)}} + \delta_o\sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}} - (\varepsilon\beta + \delta_o)t > 0. \tag{15}$$

The above α needs to satisfy the range in the proof of Proposition 2 in the Appendix A. Proposition 2 shows that when customers in a channel take the waiting time of customers in another channel as the reference point, the firm will still follow the rule of thumb for capacity planning proposed by Bassamboo et al. [34] when formulating the optimal capacity. The base capacity is $\mu_{a,1}^b = \theta\lambda_{a,s}$, $\mu_{a,2}^b = \theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}$, and the safety capacity is $\mu_{a,1}^{safe} = \sqrt{\frac{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}{c_1}}$ and $\mu_{a,2}^{safe} = \sqrt{\frac{\theta\beta(r - c_1 - c_2) + (1 - \theta)\varepsilon\beta(r - c_2)}{c_2}}$, respectively. Then, what impact will the customer’s reference behavior have on the decision and profit of the firm in this case?

4.2. Sensitivity Analysis

Lemma 4. *When offline customers regard the waiting time of online customers as their reference point, with the increase in sensitivity δ_s , the omnichannel catering firm should improve the safety capacity in the ordering stage, but not adjust the safety capacity in the production stage, i.e., $\frac{\partial \mu_{a,1}^{safe}}{\partial \delta_s} > 0$, $\frac{\partial \mu_{a,2}^{safe}}{\partial \delta_s} = 0$. When online customers take offline customers’ waiting time as the reference point, the increase in δ_o will encourage the firm to reduce the safety capacity in the ordering stage, but the safety capacity of the production stage is still unaffected, i.e., $\frac{\partial \mu_{a,1}^{safe}}{\partial \delta_o} < 0$, $\frac{\partial \mu_{a,2}^{safe}}{\partial \delta_o} = 0$.*

Lemma 4 is very different from Lemma 1. First of all, in terms of the safety capacity in the ordering stage, both online customer sensitivity δ_o and offline customer sensitivity δ_s will affect the safety capacity $\mu_{a,1}^{safe}$ in this stage. As δ_o increases, the firm should set $\mu_{a,1}^{safe}$ lower, which is very different from Lemma 1. In Lemma 1, only δ_s will affect the safety capacity, while δ_o will not. The reason for this phenomenon is that when the actual waiting time of online customers W_o is compared with that of offline customers W_s , it is actually comparing w_1 and t , so the effective shopping rate of online customers will also be affected by the waiting time of offline customers in the process of ordering food. This makes it necessary for the firm to consider δ_o when developing safety capacity in the ordering stage. Moreover, the improvement of δ_o requires the firm to reduce the safety capacity in the ordering stage, so as to increase the waiting time reference point of online customers, which is beneficial for improving the shopping rate of online customers.

The second difference between Lemma 4 and Lemma 1 is that Lemma 4 shows that the customer’s sensitivity to the reference point does not affect the safety capacity in the production stage. This is because when customers take the waiting time of customers in another channel as the reference point, customers in different channels will actually wait the same time in the production stage, and they will offset the waiting time. This allows customers to have no reference behavior to the waiting time in the production stage, so δ_o and δ_s will not affect this stage’s safety capacity.

Lemma 5. (1) *If $\frac{2\delta_o}{\beta + \delta_s} \geq \frac{\theta(r - c_1 - c_2)}{(1 - \theta)(r - c_2)}$, there must be $\frac{\partial \lambda_{a,s}}{\partial \delta_s} > 0$ and $\frac{\partial \lambda_{a,s}}{\partial \delta_o} < 0$. Once $\frac{2\delta_o}{\beta + \delta_s} < \frac{\theta(r - c_1 - c_2)}{(1 - \theta)(r - c_2)}$ and t is less than a specific value \bar{t}_1 , there must be $\frac{\partial \lambda_{a,s}}{\partial \delta_s} < 0$ and $\frac{\partial \lambda_{a,s}}{\partial \delta_o} < 0$.*
 (2) *If $\frac{\delta_o}{2(\beta + \delta_s)} \geq \frac{\theta(r - c_1 - c_2)}{(1 - \theta)(r - c_2)}$, there must be $\frac{\partial \lambda_{a,o}}{\partial \delta_o} < 0$ and $\frac{\partial \lambda_{a,o}}{\partial \delta_s} < 0$. Once $\frac{\delta_o}{2(\beta + \delta_s)} < \frac{\theta(r - c_1 - c_2)}{(1 - \theta)(r - c_2)}$ and t is less than a specific value \bar{t}_2 , there must be $\frac{\partial \lambda_{a,o}}{\partial \delta_o} > 0$ and $\frac{\partial \lambda_{a,s}}{\partial \delta_s} < 0$. (See Appendix A for proof).*

The above lemma shows that when channel information is available for the offline customers' shopping rate $\lambda_{a,s}$, if those customers' sensitivity δ_s is relatively low, i.e., $\frac{2\delta_o}{\beta+\delta_s} \geq \frac{\theta(r-c_1-c_2)}{(1-\theta)(r-c_2)}$, then the increase in δ_s can promote the improvement of the shopping rate $\lambda_{a,s}$. However, once $\frac{2\delta_o}{\beta+\delta_s} < \frac{\theta(r-c_1-c_2)}{(1-\theta)(r-c_2)}$, that is, the sensitivity δ_s is relatively high, the continuous increase in δ_s will not be conducive to the improvement of offline customers' shopping rate when the delivery time t is relatively small.

For online customers' shopping rate $\lambda_{a,o}$, when the sensitivity δ_o is high, i.e., $\frac{\delta_o}{2(\beta+\delta_s)} \geq \frac{\theta(r-c_1-c_2)}{(1-\theta)(r-c_2)}$, the continuous increase in δ_o will lead to the decrease in $\lambda_{a,o}$. In the case of smaller δ_o and t , the increase in δ_o is conducive to the increase in the online customers' shopping rate.

At the same time, we found that for a specific channel, the increase in the sensitivity of customers in different channels to the reference point led to the decrease in customers' shopping rate in this specific channel. This cross-channel effect is significant when channel information is available. We take the offline channel customers' increasing sensitivity to reference points as an example. When δ_s increases, it indicates that the sensitivity of offline channel customers to the reference point increases, and the firm should increase W_o so as to bring higher utility to offline customers. However, the increase in W_o leads to the decrease in the online customers' shopping rate. This cross-channel influence is also an issue that omnichannel catering firms need to pay attention to.

According to Proposition 2, we obtain the profit of the firm as follows:

$$\pi_a = \theta(\alpha + \delta_s t) \frac{(r - c_1 - c_2) + (1 - \theta)(\alpha - (\varepsilon\beta + \delta_o)t)(r - c_2)}{-2\sqrt{c_1(\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2))} - 2\sqrt{c_2(\theta\beta(r - c_1 - c_2) + (1 - \theta)\varepsilon\beta(r - c_2))}}$$

Lemma 6. When the actual waiting time of offline customers in the system satisfies $w_1 > t$, the profit of the firm decreases with the rise in δ_s , but increases with the rise in δ_o , i.e., $\frac{\partial\pi_a}{\partial\delta_s} < 0$ and $\frac{\partial\pi_a}{\partial\delta_o} > 0$. In contrast, the profit increases with the increase in δ_s and decreases with the increase in δ_o , i.e., $\frac{\partial\pi_a}{\partial\delta_s} > 0$ and $\frac{\partial\pi_a}{\partial\delta_o} < 0$. (See Appendix A for proof).

Once customers in different channels take the waiting time of customers in another channel as the reference point, the change of the optimal profit is affected by the size relationship between w_1 and t . This is different from Lemma 3, because once customers have the reference behavior and take the waiting time of customers in different channels as their reference points, then the difference between offline customers and online customers lies in the waiting time for the ordering process and the waiting time for the delivery process. Since the delivery is carried out by a third-party platform, the delivery time t is an exogenous variable. Then, the only thing the firm can decide is the capacity in the ordering stage. We find that the waiting time in the ordering stage can be expressed as follows:

$$w_1 = \sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}}$$

It can be seen that when the delivery time t is a fixed value, and lower c_1 and c_2 can effectively reduce the waiting time of the ordering stage. However, c_1 and c_2 are actually the wages of front desk and kitchen staff in the labor market, which are difficult to change in a short period of time in a perfectly competitive labor market. The proportion θ of offline customers is determined by the market segment. Therefore, for the firm, the only means to reduce w_1 is to try its best to reduce the customer's unit waiting cost β , and the sensitivity δ_s and δ_o to the reference point.

4.3. Comparison of Different Situations

In the previous section, we analyzed the omnichannel catering firm’s optimal decisions and profit based on the reference behavior under the conditions of different information availability. In this section, we compare the optimal decisions and profit when channel information is available or not.

Lemma 7. *Under the condition that channel information is available, the safety capacity in different stages is smaller than that of the corresponding stages under the condition that channel information is unavailable, i.e., $\mu_{a,1}^{safe} < \mu_1^{safe}$ and $\mu_{a,2}^{safe} < \mu_2^{safe}$.*

The above lemma shows that the safety capacity at each stage will decrease in the case where information in different channels is available compared with that in the case where information is unavailable, which is conducive to the reduction of cost.

This is because the customer’s reference behavior is actually more rational when channel information is available. Through further exploration, we find that customers only refer to the ordering time and delivery time when information is available. However, once information is not available, their waiting time for reference also includes the waiting time in the food production stage, which will greatly amplify this reference behavior. The expansion of the impact of customer reference behavior actually puts forward higher requirements for the capacities of the firm.

Meanwhile, we can obtain the management implications from the above lemma; that is, for the operation of the firm, enhancing the transparency and availability of information between channels would be conducive to reducing the safety capacity, so as to reduce the cost. The way to enhance information transparency is to publicize the offline customer’s waiting time in online channels and the queue length of the different channels in the offline channels. This practice is currently used in KFC and McDonald’s. In these firms, we can often see the length of the queue on the bulletin board after we order, and this also includes the length of the queue in the online channel.

In the following, we compare the differences in profits between different cases. We can decompose the firm’s profit under the condition of information availability as follows:

$$\pi_a = R_{a1} + R_{a2} - C_{a1} - C_{a2}, \tag{16}$$

where $R_{a1} = \frac{\theta(\alpha + \delta_s t)(r - c_1 - c_2)}{\sqrt{c_1(\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2))}}$, $R_{a2} = \frac{(1 - \theta)(\alpha - (\epsilon\beta + \delta_o)t)(r - c_2)}{2\sqrt{c_2(\theta\beta(r - c_1 - c_2) + (1 - \theta)\epsilon\beta(r - c_2))}}$, $C_{a1} = 2$

The above formula does not mean that profit is independent of capacity, but that capacity is implied in the formula of R_{a1} , R_{a2} , C_{a1} and C_{a2} . Similarly, the firm’s profit in the case of unavailable information can also be decomposed, i.e.,

$$\pi_n = R_{n1} + R_{n2} - C_{n1} - C_{n2}, \tag{17}$$

where $R_{n1} = \frac{\theta(\alpha + \delta_s \bar{W}_s)(r - c_1 - c_2)}{\sqrt{c_1\theta(\beta + \delta_s)(r - c_1 - c_2)}}$, $R_{n2} = \frac{(1 - \theta)(\alpha - (\epsilon\beta + \delta_o)t + \delta_o \bar{W}_o)(r - c_2)}{2\sqrt{c_2(\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\epsilon\beta + \delta_o)(r - c_2))}}$, $C_{n1} = 2$

It can be seen from the above that the profit under different cases can be divided into four parts for analysis. Terms R_{n1} and R_{a1} in the above profit formula represent the revenue obtained from offline customers, terms R_{n2} and R_{a2} represent the revenue obtained from online customers, terms C_{n1} and C_{a1} represent costs that the firm need to bear in the ordering stage, and terms C_{n2} and C_{a2} represent costs that the firm need to bear in the production stage. According to Lemma 7, the following Lemma 8 can be obtained directly.

Lemma 8. *When information is available, the omnichannel catering firm should bear less costs, i.e., $C_{a1} + C_{a2} < C_{n1} + C_{n2}$.*

This indicates that improving the availability of channel information will help reduce costs for the firm. The benefits brought by such information transparency and availability

are one of the reasons why firms attach importance to information flow at present. In the analysis of Lemma 7, we also gave the methods and suggestions to increase the availability of information between channels, which is beneficial for the development of omnichannel catering firms.

By comparing R_{a1} and R_{n1} , it can be seen that if $t > \bar{W}_s$, the firm obtains more revenue from offline customers when information is available. However, from the comparison of R_{a2} and R_{n2} , it can be seen that when information is unavailable, the firm obtains more revenue from online customers. Thus, we obtain the following lemma.

Lemma 9. *The size relationship between π_a and π_n is still affected by the size of \bar{W}_s and \bar{W}_o . If \bar{W}_s and \bar{W}_o are small, then $\pi_a > \pi_n$. However, once \bar{W}_s and \bar{W}_o are both relatively large, the firm can obtain greater profits from customers in online and offline channels when channel information is unavailable, and thus $\pi_a < \pi_n$.*

In actual operation, the firm can directly compare the profits between different cases to choose more favorable information disclosure decisions. However, for omnichannel catering firms, it is obviously beneficial to enhance the availability of information between channels. According to Lemma 7 and 8, we know that the availability of channel information can reduce cost expenditure by changing customers' reference behavior, which is of great significance for firms to reduce costs. Of course, this is not static. When the customer's expected waiting time is high, it is obviously beneficial to reduce the availability of channel information. In the next section, numerical analysis is given to further illustrate this point.

5. Numerical Analysis

In the previous section, based on the theoretical model, we analyzed how the omnichannel catering firm makes decisions on optimal capacities based on customer reference behaviors in different channels, and the impact of customer sensitivity to reference points on optimal capacities and profits. In order to dynamically grasp the impact of customer sensitivity changes further in the different channels, we conducted further research through numerical analysis in this section.

5.1. Analysis of Offline Customers

Our parameter setting was based on the research of Yuan et al. [3] and Zhan et al. [5], so that relevant firms can select the appropriate values for analysis according to their specific conditions in operation and the results of market research, which makes our research more practical. This application is also shown at the end of this section. In this market, we assume that the proportion of offline customers occupying the market is $\theta = 0.6$, and the customer's sensitivity to waiting time is $\beta = 0.5$. Among them, the sensitivity of online customers to the reference point is $\delta_o = 0.5$, and the proportion of their anxiety cost is $\varepsilon = 0.3$. In addition, these online customers also need to go through a delivery process with time $t = 0.5$. In this case, the base shopping rate of online and offline customers is $\alpha = 5$, and the expectation of waiting time is $\bar{W}_s = \bar{W}_o = 1$. However, for the firm, the revenue obtained from each customer is $r = 2$, and the unit cost to be borne at each stage is $c_1 = 0.2$, $c_2 = 0.3$. As δ_s changes, the firm's optimal capacity, profit, and the effective shopping rate of customers is shown in Figure 1.

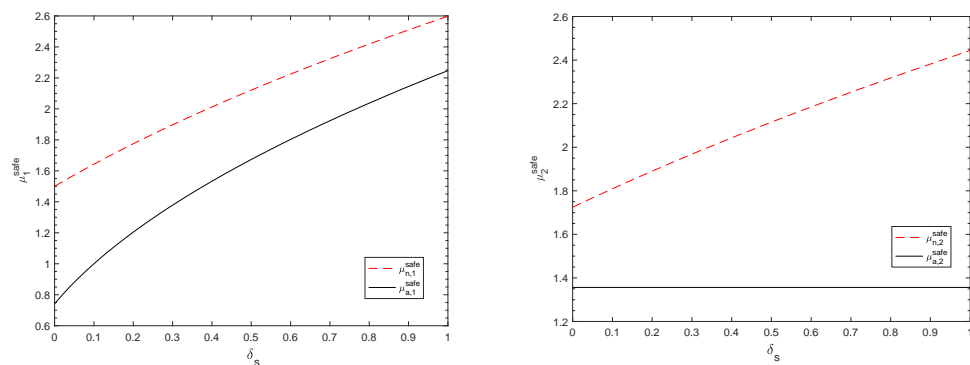


Figure 1. Left panel illustrates the safety capacity μ_1^{safe} in the ordering stage with respect to δ_s , and right panel illustrates the safety capacity μ_2^{safe} in the production stage with respect to δ_s .

The above image firstly shows that with the increase in offline customers’ sensitivity to reference points, the safety capacity set by omnichannel catering firms does not decrease, and the safety capacities of all stages in the case of unavailable information is higher than that in the case of available information. Especially in the safety capacity of the production stage, with the increase in δ_s , the difference of the safety capacity between the two cases gradually expands.

Figure 2 shows an interesting phenomenon; that is, with the change of δ_s , the offline customer’s shopping rate $\lambda_{n,s}$ in the case of unavailable information will always be higher than $\lambda_{a,s}$ in the case of available information. Moreover, with the increase in δ_s , the shopping rate of offline customers will increase whether the channel information is available or not available. This is because the increase in δ_s will promote the firm to improve the safety capacity according to Lemma 1 and Lemma 4, thus reducing the waiting time of customers in the system and increasing their utility, which leads to the increase in customers’ shopping rate.

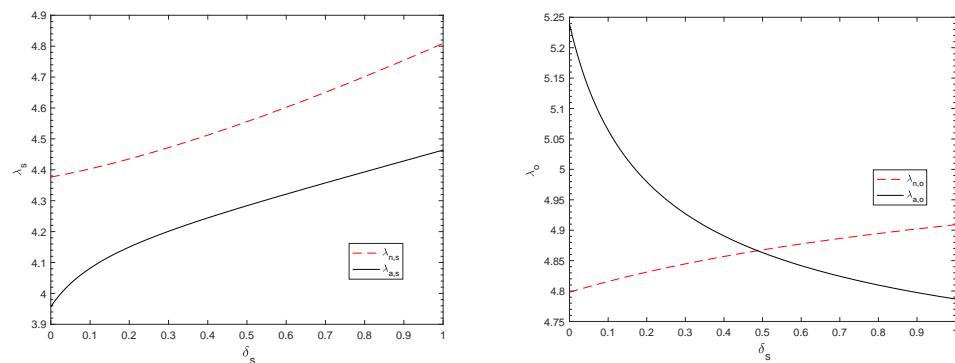


Figure 2. Left panel illustrates the offline customers’ shopping rate λ_s with respect to δ_s , and right panel illustrates the online customers’ shopping rate λ_o with respect to δ_s .

However, not so for online customers’ shopping rate λ_o . The increase in δ_s will promote the increase in $\lambda_{n,o}$ and a decrease in $\lambda_{a,o}$, but $\lambda_{a,o}$ is significantly higher than $\lambda_{n,o}$ when $\delta_s = 0$. Thus, as the figure shows, an increase in δ_s will cause the gap between $\lambda_{n,o}$ and $\lambda_{a,o}$ to narrow and then widen, and this change will eventually show up in the firm’s profits.

As shown in Figure 3, with the change in δ_s , the optimal profit π_a in the case with information available is always higher than π_n in the case with information unavailable. Although offline customers have a higher shopping rate when information is unavailable, the cost savings result in higher profits due to lower safety capacity at each stage when information is available. Meanwhile, the difference in profit will be reduced, because the

online customers' shopping rate $\lambda_{a,o}$ decreases gradually with the increase in δ_s , which makes π_a tend to π_n .

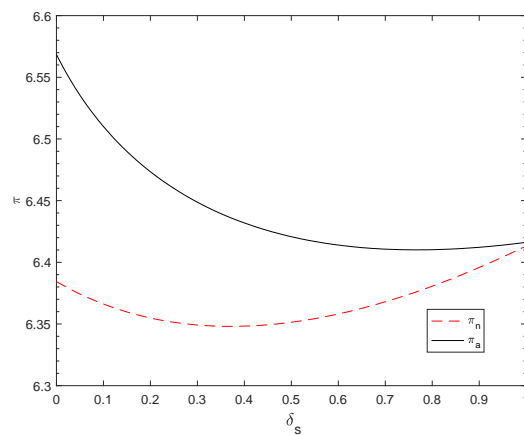


Figure 3. The firm's profit with respect to δ_s .

5.2. Analysis of Online Customers

In this section, we conduct a numerical analysis of the impact of online customer sensitivity δ_o . We assume that the offline customer's sensitivity to the reference point is $\delta_s = 0.5$, and with the change in online customer's sensitivity δ_o , the firm's optimal capacity, profit and customer's effective shopping rate will be shown in the following figure.

Similar to Figure 1, we still find that the safety capacity of each stage in the case with information unavailable is still higher than that in the case with information available in Figure 4. Meanwhile, we can see from the figure that with the increase in δ_o , the difference in safety capacity between the two cases will gradually increase. In the left panel, $\mu_{n,1}^{safe}$ is fixed, and as $\mu_{a,1}^{safe}$ decreases, the difference between them widens. However, in the right panel, $\mu_{a,2}^{safe}$ is fixed, and as $\mu_{n,2}^{safe}$ increases, the difference between them widens.

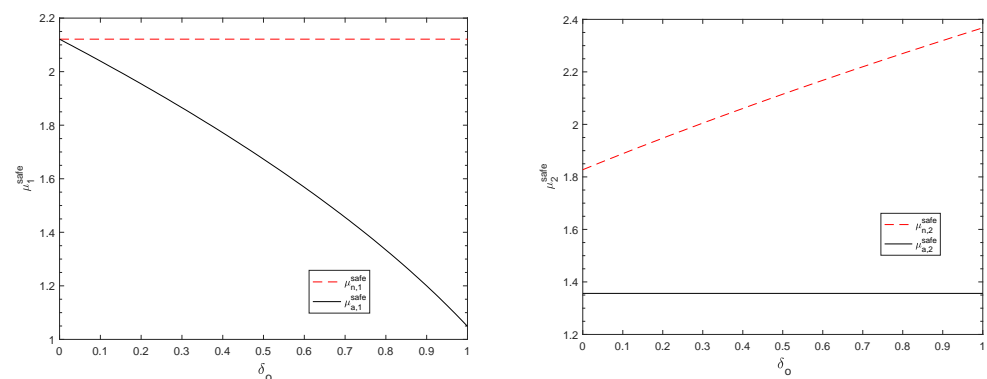


Figure 4. Left panel illustrates the safety capacity $\mu_{n,1}^{safe}$ in ordering stage with respect to δ_o , and right panel illustrates the safety capacity $\mu_{n,2}^{safe}$ in production stage with respect to δ_o .

From the left panel of Figure 5, we can see that the difference between the offline customers' shopping rate in the cases with and without information availability will gradually expand with the decrease in $\lambda_{a,s}$ and the increase in $\lambda_{n,s}$, which indicates that for offline customers' shopping rate, information unavailability is beneficial. This is because with the increase in δ_o , the firm will increase the safety capacity in the production stage to reduce the waiting time of customers in the case of information unavailability according to Lemma 1, which makes the customer shopping rate, in this case, continue to increase. According to Lemma 4, in the case of information availability, the firm will reduce the safety capacity in the ordering stage, which leads to an increase in the waiting time of offline

customers and a decline in the shopping rate. Thus, the increase in $\lambda_{n,s}$ and the decrease in $\lambda_{a,s}$ lead to the greater difference between the two.

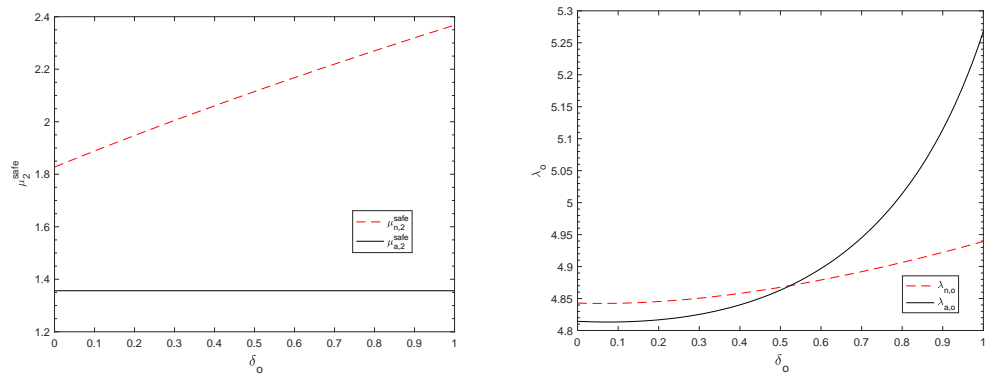


Figure 5. Left panel illustrates the offline customers’ shopping rate λ_s with respect to δ_o , and right panel illustrates the online customers’ shopping rate λ_o with respect to δ_o .

In contrast, with the increase in δ_o , the shopping rate of online customers will increase no matter whether the channel information is available or not, because it can be found from Lemma 1 and Lemma 4 that the increase in δ_o prompts the firm to improve the safety capacity, thus reducing the waiting time of customers and increasing their shopping rate. Meanwhile, the right panel of Figure 5 shows that the availability of channel information leads to a higher shopping rate for online customers as δ_o increases. Therefore, for customers from different channels, the availability of channel information has different effects.

Figure 6 clearly shows that when channel information is available, the omnichannel catering firm can obtain higher profits through online customers’ reference behavior. The rise in profit can be explained in two ways. On the one hand, based on Figure 4, it can be seen that when information is available, the firm can develop lower safety capacities, which can effectively reduce the cost input. Secondly, Figure 5 shows that although the increase in δ_o reduces offline customers’ shopping rate, it can make online customers’ shopping rate increase sharply. Therefore, combined with Figure 3, we find that the transparency of channel information and the existence of customer reference behavior are of great significance to the improvement of the firm’s profit. Moreover, since people in different countries and regions have different consumption preferences and risk aversion [35], omnichannel firms should first conduct market research on the characteristics of customers in the market when making corresponding optimal decisions, especially customers’ reference points and corresponding sensitivities. Only in this way can firms make decisions more in line with the market demand to obtain the optimal profits.

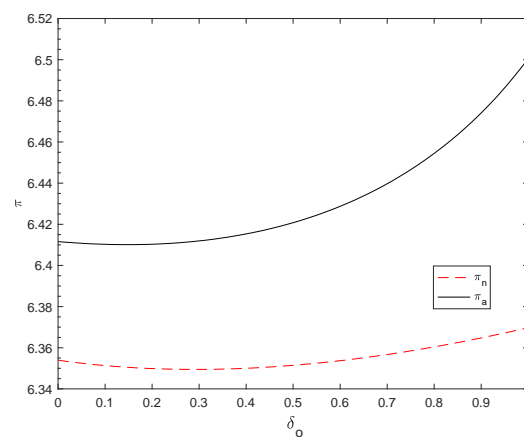


Figure 6. The firm’s profit with respect to δ_o .

Let us take an example to illustrate the application of the findings in this article. We took the operation of catering firm A as an example. Firm A operates both online channels and offline channels. After market research, it is known that the proportion of offline customers in the market $\theta = 0.6$ and the sensitivity of reference point $\delta_s = 0.5$. The sensitivity of online customers to the reference point is $\delta_o = 0.6$, and they need to experience a delivery process with time $t = 0.5$. Meanwhile, the base shopping rate of both online and offline customers is $\alpha = 5$, which is derived from the quality and taste of the food. These customers' expected waiting time is $\bar{W}_s = \bar{W}_o = 1$, sensitivity to waiting time is $\beta = 0.5$. For this firm A, the revenue received from each customer is $r = 2$, and the average cost of employing the receptionist to serve each customer is $c_1 = 0.2$, while the average cost of employing the chef to produce for each customer is $c_2 = 0.3$. According to the calculation of profit, it can be obtained that the profit under the condition that information is unavailable is $\pi_n = 6.35$, while the profit under the condition that information is available is $\pi_a = 6.43$, obviously $\pi_a > \pi_n$. Therefore, at this time, the firm should publish the waiting information of customers in various channels, which can help firm A gain higher profits. The above application for decision-making reflects the practical value of our paper.

6. Conclusions

We analyzed the optimal capacity decision of the omnichannel catering firm based on the reference behavior of customers in offline and online channels. Firstly, according to the availability of channel information, the operation situation of firms was divided into the situation where channel information was unavailable and the situation where channel information was available. Then, we determined the optimal capacity decisions of firms in different situations and carried out further study through numerical analysis. Based on the solution of the theoretical model of optimal decision and numerical analysis, the following conclusions were obtained:

- (1) If channel information is unavailable, customers take their expectations of waiting time as the reference point. With the improvement of the reference point sensitivity of customers, the omnichannel catering firm should enhance the safety capacity of the ordering and production stages to meet the needs of customers. At the same time, the improvement of the reference point sensitivity of customers also contributes to the increase in the firm's profit under the higher customer expected waiting time.
- (2) Once channel information is available, customers take the waiting time of customers in other channels as the reference point. In this case, the change of customer sensitivity does not affect the safety capacity in the production stage. At this time, the sensitivity of online customers and the sensitivity of offline customers have exactly opposite effects on the safety capacity in the ordering stage. Moreover, the impact of customer sensitivity on the firm's profit is also different between the different channels.
- (3) We also found that when channel information is available, the firm can set lower safety capacities to reduce costs, and can obtain higher profit when customers' expected waiting time is small than when information is unavailable. In fact, this provides theoretical support for the choice of the channel information disclosure strategy of firms, and is of great significance for the operation and development of catering enterprises.

In terms of research significance, this paper uses the queuing theory to build an optimal capacity decision model of omnichannel catering firms based on the reference behavior of customers. This paper solves the problem of the capacity formulation of omnichannel catering firms under the conditions of both information channel availability and non-availability, which has important practical significance for the development of such firms. In addition, we introduce reference theory into the capability management of firms, which expands the scope of the application of reference theory, and enriches capacity management theory at the same time.

This paper does not analyze the loss aversion behavior of customers, and the loss aversion behavior of customers in different channels will affect their final shopping rate, which may have an impact on firms. Therefore, in subsequent research, we will study the

loss aversion behavior of customers in different channels to further analyze the impact of customer behavior on omnichannel catering firms.

Author Contributions: Conceptualization, W.-T.Z.; methodology, W.-T.Z.; software, X.-P.W., H.J. and D.H.; validation, W.-T.Z., M.-H.J. and X.-P.W.; formal analysis, W.-T.Z. and Y.-T.L.; investigation, W.-T.Z., D.H. and X.-P.W.; resources, M.-H.J.; writing—original draft preparation, W.-T.Z.; writing—review and editing, W.-T.Z., M.-H.J., X.-P.W., H.J., D.H. and Y.-T.L.; supervision, M.-H.J.; project administration, M.-H.J.; funding acquisition, M.-H.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the National Natural Science Foundation of China Grant No. 71831005 and No. 71502044.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Proposition 1. Omnichannel caterers will choose appropriate optimal capacities $\mu_{n,1}$ and $\mu_{n,2}$ to maximize profit in the case of unavailable information, i.e.,

$$\begin{aligned} & \max_{\theta\lambda_s < \mu_1, \theta\lambda_s + (1-\theta)\lambda_o < \mu_2} \{r(\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o}) - c_1\mu_{n,1} - c_2\mu_{n,2}\} \\ \text{s.t. } & \lambda_{n,s} = [\alpha - \beta(w_1(\mu_{n,1}, \theta\lambda_{n,s}) + w_2(\mu_{n,2}, (\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o}))) - \delta_s(W_s - \bar{W}_s)]^+ \\ & \lambda_{n,o} = [\alpha - \varepsilon\beta(w_2(\mu_{n,2}, (\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o})) + t) - \delta_o(W_o - \bar{W}_o)]^+ \end{aligned}$$

It can be seen from Table A1 that $w_1(\mu_{n,1}, \theta\lambda_{n,s})$ and $w_2(\mu_{n,2}, (\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o}))$ in the above equation are the waiting times of customers in the first and second stages, respectively. According to the calculation formula of waiting time in M/M/1 queuing model, we find that $w_1(\mu_{n,1}, \theta\lambda_{n,s}) = \frac{1}{\mu_{n,1} - \theta\lambda_{n,s}}$, $w_2(\mu_{n,2}, (\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o})) = \frac{1}{\mu_{n,2} - (\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o})}$. Meanwhile, according to the hypothesis behind Formula (1) in the text, we can see that $W_s = w_1 + w_2$, $W_o = w_2 + t$ are the total waiting times of offline customers and online customers in the system, respectively. These two formulas can be substituted into the above optimization model, and thus $\delta_s W_s = \delta_s(w_1 + w_2)$, $\delta_o W_o = \delta_o(w_2 + t)$. Therefore, the above optimization model can be written as follows:

$$\begin{aligned} & \max_{\theta\lambda_s < \mu_1, \theta\lambda_s + (1-\theta)\lambda_o < \mu_2} \{r(\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o}) - c_1\mu_{n,1} - c_2\mu_{n,2}\} \\ \text{s.t. } & \lambda_{n,s} = [\alpha - (\beta + \delta_s)(w_1(\mu_{n,1}, \theta\lambda_{n,s}) + w_2(\mu_{n,2}, (\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o}))) + \delta_s \bar{W}_s]^+ \\ & \lambda_{n,o} = [\alpha - (\varepsilon\beta + \delta_o)(w_2(\mu_{n,2}, (\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o})) + t) + \delta_o \bar{W}_o]^+ \end{aligned}$$

The above equation shows that $\delta_s W_s$ and $\delta_o W_o$ are replaced by $\delta_s(w_1 + w_2)$ and $\delta_o(w_2 + t)$, respectively. Since $w_1(\mu_{n,1}, \theta\lambda_{n,s}) = \frac{1}{\mu_{n,1} - \theta\lambda_{n,s}}$ and $w_2(\mu_{n,2}, (\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o})) = \frac{1}{\mu_{n,2} - (\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o})}$, the actual variables in the optimization model are $\lambda_{n,s}$, $\lambda_{n,o}$, $\mu_{n,1}$, $\mu_{n,2}$. We use the Lagrange multiplier method to solve the above optimal model. Then, the Lagrangian function of the above formula is defined as follows:

$$\begin{aligned} L(\lambda_{n,s}, \lambda_{n,o}, \mu_{n,1}, \mu_{n,2}, M, N) &= r(\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o}) - c_1\mu_{n,1} - c_2\mu_{n,2} \\ &+ M \left(\lambda_{n,s} - \alpha + (\beta + \delta_s) \frac{1}{\mu_{n,1} - \theta\lambda_{n,s}} + (\beta + \delta_s) \frac{1}{\mu_{n,2} - (\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o})} - \delta_s \bar{W}_s \right) \\ &+ N \left(\lambda_{n,o} - \alpha + (\varepsilon\beta + \delta_o) \frac{1}{\mu_{n,2} - (\theta\lambda_{n,s} + (1-\theta)\lambda_{n,o})} + (\varepsilon\beta + \delta_o)t - \delta_o \bar{W}_o \right) \end{aligned}$$

Table A1. Notations and definitions.

Notations	Definitions
$\mu_i, i = 1, 2$	Capacity in stage $i = 1, 2$
$\mu_i^b, i = 1, 2$	Base capacity in stage $i = 1, 2$
$\mu_i^{safe}, i = 1, 2$	Safety capacity in stage $i = 1, 2$
$c_i, i = 1, 2$	Unit capacity cost in stage $i = 1, 2$
$w_i, i = 1, 2$	Waiting time in stage $i = 1, 2$
r	The revenue per customer
θ	The proportion of offline customers in the market
λ_s or λ_o	The actual shopping rate of offline or online customers
Λ	Total market demand
\bar{W}_s or \bar{W}_o	The expected waiting times (reference point) of offline or online customers
t	The delivery time to online customers
δ_s or δ_o	The sensitivity of offline or online customers to reference points
β	Customer sensitivity to waiting time
ε	The proportion of anxiety cost in waiting cost per unit time
α	The base shopping rate of online and offline customers (determined by food quality)
π	The firm's profits

M and N in the above equation are the parameters introduced when using the Lagrange multiplier method. Because $\delta_s W_s$ and $\delta_o W_o$ are replaced by $\delta_s(w_1 + w_2)$ and $\delta_o(w_2 + t)$, respectively, and $w_1(\mu_{n,1}, \theta\lambda_{n,s}) = \frac{1}{\mu_{n,1} - \theta\lambda_{n,s}}$, $w_2(\mu_{n,2}, (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o})) = \frac{1}{\mu_{n,2} - (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o})}$, then $\delta_s W_s$ and $\delta_o W_o$ are not shown in $L(\lambda_{n,s}, \lambda_{n,o}, \mu_{n,1}, \mu_{n,2}, M, N)$. Thus, we obtain:

$$\frac{\partial L}{\partial \lambda_{n,s}} = r\theta + M + \frac{M\theta(\beta + \delta_s)}{(\mu_{n,1} - \theta\lambda_{n,s})^2} + \frac{M\theta(\beta + \delta_s)}{(\mu_{n,2} - (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o}))^2} + \frac{N\theta(\varepsilon\beta + \delta_o)}{(\mu_{n,2} - (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o}))^2} = 0$$

$$\frac{\partial L}{\partial \lambda_{n,o}} = r(1 - \theta) + N + \frac{M(\beta + \delta_s)(1 - \theta)}{(\mu_{n,2} - (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o}))^2} + \frac{N(\varepsilon\beta + \delta_o)(1 - \theta)}{(\mu_{n,2} - (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o}))^2} = 0$$

$$\frac{\partial L}{\partial \mu_{n,1}} = -c_1 - \frac{M(\beta + \delta_s)}{(\mu_{n,1} - \theta\lambda_{n,s})^2} = 0$$

$$\frac{\partial L}{\partial \mu_{n,2}} = -c_2 - \frac{M(\beta + \delta_s)}{(\mu_{n,2} - (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o}))^2} - \frac{N(\varepsilon\beta + \delta_o)}{(\mu_{n,2} - (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o}))^2} = 0$$

$$\frac{\partial L}{\partial M} = \lambda_{n,s} - \alpha + (\beta + \delta_s) \frac{1}{\mu_{n,1} - \theta\lambda_{n,s}} + (\beta + \delta_s) \frac{1}{\mu_{n,2} - (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o})} - \delta_s \bar{W}_s = 0$$

$$\frac{\partial L}{\partial N} = \lambda_{n,o} - \alpha + (\varepsilon\beta + \delta_o) \frac{1}{\mu_{n,2} - (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o})} + (\varepsilon\beta + \delta_o)t - \delta_o \bar{W}_o = 0$$

$c_1 = -\frac{M(\beta + \delta_s)}{(\mu_{n,1} - \theta\lambda_{n,s})^2}$ and $c_2 = -\frac{M(\beta + \delta_s)}{(\mu_{n,2} - (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o}))^2} - \frac{N(\varepsilon\beta + \delta_o)}{(\mu_{n,2} - (\theta\lambda_{n,s} + (1 - \theta)\lambda_{n,o}))^2}$ can be obtained from the above formulas $\frac{\partial L}{\partial \mu_{n,1}} = 0$ and $\frac{\partial L}{\partial \mu_{n,2}} = 0$, and put them into $\frac{\partial L}{\partial \lambda_{n,s}} = 0$ and $\frac{\partial L}{\partial \lambda_{n,o}} = 0$, then $M = \theta(r - c_1 - c_2)$ and $N = (1 - \theta)(r - c_2)$ can be obtained. Therefore, the optimal capacity of each stage can be obtained by $\frac{\partial L}{\partial \mu_{n,1}} = 0$ and $\frac{\partial L}{\partial \mu_{n,2}} = 0$ as follows:

$$\mu_{n,1} = \theta\lambda_{n,s} + \sqrt{\frac{\theta(\beta + \delta_s)(r - c_1 - c_2)}{c_1}}$$

$$\mu_{n,2} = \theta\lambda_{n,s} + (1 - \theta)\lambda_o + \sqrt{\frac{\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\varepsilon\beta + \delta_o)(r - c_2)}{c_2}}$$

Then plug the above optimal capabilities into $\frac{\partial L}{\partial M} = 0$ and $\frac{\partial L}{\partial N} = 0$ to obtain the following customer shopping rate function:

$$\lambda_{n,s} = \alpha - (\beta + \delta_s) \sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2)}} - (\beta + \delta_s) \sqrt{\frac{c_2}{\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\epsilon\beta + \delta_o)(r - c_2)}} + \delta_s \bar{W}_s,$$

$$\lambda_{n,o} = \alpha - (\epsilon\beta + \delta_o) \sqrt{\frac{c_2}{\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\epsilon\beta + \delta_o)(r - c_2)}} - (\epsilon\beta + \delta_o)t + \delta_o \bar{W}_o.$$

According to the research of Gao et al. [4] and Zhan et al. [5], the above effective shopping rates $\lambda_{n,s}$ and $\lambda_{n,o}$ must be positive in actual operation. Therefore, according to $\lambda_{n,s} > 0$ and $\lambda_{n,o} > 0$, the effective range of the base shopping rate α with practical significance can be obtained as follows:

$$\begin{cases} \alpha > (\beta + \delta_s) \sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2)}} + (\beta + \delta_s) \sqrt{\frac{c_2}{\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\epsilon\beta + \delta_o)(r - c_2)}} - \delta_s \bar{W}_s, \\ \alpha > (\epsilon\beta + \delta_o) \sqrt{\frac{c_2}{\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\epsilon\beta + \delta_o)(r - c_2)}} + (\epsilon\beta + \delta_o)t - \delta_o \bar{W}_o. \end{cases}$$

Therefore, when the base shopping rate α meets the above formula, the effective shopping rate of customers is positive and has significance, i.e., $\lambda_{n,s} > 0$ and $\lambda_{n,o} > 0$. However, if α does not meet the above formula, it does not meet the practical significance and is not included in the research scope of this paper. We also make a corresponding explanation at the end of Proposition 1. □

Proof of Proposition 2. Omnichannel caterers will choose appropriate optimal capacities $\mu_{a,1}$ and $\mu_{a,2}$ to maximize profit in the case of available information, i.e.,

$$\max_{\theta\lambda_{a,s} < \mu_{a,1}, \theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o} < \mu_{a,2}} \{r(\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}) - c_1\mu_{a,1} - c_2\mu_{a,2}\}$$

$$\begin{aligned} \text{s.t. } \lambda_{a,s} &= [\alpha - \beta(w_1(\mu_{a,1}, \theta\lambda_{a,s}) + w_2(\mu_{a,2}, (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}))) - \delta_s(W_s - W_o)]^+ \\ \lambda_{a,o} &= [\alpha - \epsilon\beta(w_2(\mu_{a,2}, (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}))) + t - \delta_o(W_o - W_s)]^+ \end{aligned}$$

It can be seen from Table A1 that $w_1(\mu_{a,1}, \theta\lambda_{a,s})$ and $w_2(\mu_{a,2}, (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}))$ in the above equation are the waiting times of customers in the first and second stages, respectively. According to the calculation formula of waiting time in the M/M/1 queuing model, we find that $w_1(\mu_{a,1}, \theta\lambda_{a,s}) = \frac{1}{\mu_{a,1} - \theta\lambda_{a,s}}$, $w_2(\mu_{a,2}, (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})) = \frac{1}{\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})}$. Meanwhile, we know that $W_s = w_1 + w_2$, $W_o = w_2 + t$ are the total waiting time of offline customers and online customers in the system, respectively. These two formulas can be substituted into the above optimization model; therefore, the above optimization model can be written as follows:

$$\max_{\theta\lambda_{a,s} < \mu_{a,1}, \theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o} < \mu_{a,2}} \{r(\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}) - c_1\mu_{a,1} - c_2\mu_{a,2}\}$$

$$\begin{aligned} \text{s.t. } \lambda_{a,s} &= [\alpha - (\beta + \delta_s)w_1(\mu_{a,1}, \theta\lambda_{a,s}) - \beta w_2(\mu_{a,2}, (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})) + \delta_s t]^+ \\ \lambda_{a,o} &= [\alpha - \epsilon\beta w_2(\mu_{a,2}, (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})) - (\epsilon\beta + \delta_o)t + \delta_o w_1(\mu_{a,1}, \theta\lambda_{a,s})]^+ \end{aligned}$$

and since $w_1(\mu_{a,1}, \theta\lambda_{a,s}) = \frac{1}{\mu_{a,1} - \theta\lambda_{a,s}}$ and $w_2(\mu_{a,2}, (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})) = \frac{1}{\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})}$, the actual variables in the optimization model are $\lambda_{a,s}$, $\lambda_{a,o}$, $\mu_{a,1}$, $\mu_{a,2}$. We use the Lagrange multiplier method to solve the above optimal model. Then, the Lagrangian function of the above formula is defined as follows:

$$\begin{aligned} L(\lambda_{a,s}, \lambda_{a,o}, \mu_{a,1}, \mu_{a,2}, M, N) &= r(\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}) - c_1\mu_{a,1} - c_2\mu_{a,2} \\ &+ M \left(\lambda_{a,s} - \alpha + (\beta + \delta_s) \frac{1}{\mu_{a,1} - \theta\lambda_{a,s}} + \beta \frac{1}{\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})} - \delta_s t \right) \\ &+ N \left(\lambda_{a,o} - \alpha + \epsilon\beta \frac{1}{\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})} + (\epsilon\beta + \delta_o)t - \delta_o \frac{1}{\mu_{a,1} - \theta\lambda_{a,s}} \right) \end{aligned}$$

M and N in the above equation are the parameters introduced when using the Lagrange multiplier method. Because W_s and W_o are replaced by $w_1 + w_2$ and $w_2 + t$, respectively, and $w_1(\mu_{a,1}, \theta\lambda_{a,s}) = \frac{1}{\mu_{a,1} - \theta\lambda_{a,s}}$, $w_2(\mu_{a,2}, (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})) = \frac{1}{\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})}$, then W_s and W_o are not shown in $L(\lambda_{a,s}, \lambda_{a,o}, \mu_{a,1}, \mu_{a,2}, M, N)$. Thus, we can obtain:

$$\begin{aligned} \frac{\partial L}{\partial \lambda_{a,s}} &= r\theta + M + \frac{M\theta(\beta + \delta_s)}{(\mu_{a,1} - \theta\lambda_{a,s})^2} + \frac{M\theta\beta}{(\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}))^2} + \frac{N\theta\epsilon\beta}{(\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}))^2} \\ &\quad - \frac{N\theta\delta_o}{(\mu_{a,1} - \theta\lambda_{a,s})^2} = 0 \\ \frac{\partial L}{\partial \lambda_{a,o}} &= r(1 - \theta) + N + \frac{M\beta(1 - \theta)}{(\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}))^2} + \frac{N\epsilon\beta(1 - \theta)}{(\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}))^2} = 0 \\ \frac{\partial L}{\partial \mu_{a,1}} &= -c_1 - \frac{M(\beta + \delta_s)}{(\mu_{a,1} - \theta\lambda_{a,s})^2} + \frac{N\delta_o}{(\mu_{a,1} - \theta\lambda_{a,s})^2} = 0 \\ \frac{\partial L}{\partial \mu_{a,2}} &= -c_2 - \frac{M\beta}{(\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}))^2} - \frac{N\epsilon\beta}{(\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}))^2} = 0 \\ \frac{\partial L}{\partial M} &= \lambda_{a,s} - \alpha + (\beta + \delta_s)\frac{1}{\mu_{a,1} - \theta\lambda_{a,s}} + \beta\frac{1}{\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})} - \delta_s t = 0 \\ \frac{\partial L}{\partial N} &= \lambda_{a,o} - \alpha + \epsilon\beta\frac{1}{\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o})} + (\epsilon\beta + \delta_o)t - \delta_o\frac{1}{\mu_{a,1} - \theta\lambda_{a,s}} = 0 \\ c_1 &= -\frac{M(\beta + \delta_s)}{(\mu_{a,1} - \theta\lambda_{a,s})^2} + \frac{N\delta_o}{(\mu_{a,1} - \theta\lambda_{a,s})^2} \text{ and } c_2 = -\frac{M\beta}{(\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}))^2} - \frac{N\epsilon\beta}{(\mu_{a,2} - (\theta\lambda_{a,s} + (1 - \theta)\lambda_{a,o}))^2} \end{aligned}$$

can be obtained from the above formulas $\frac{\partial L}{\partial \mu_{a,1}} = 0$ and $\frac{\partial L}{\partial \mu_{a,2}} = 0$, and put them into $\frac{\partial L}{\partial \lambda_{a,s}} = 0$ and $\frac{\partial L}{\partial \lambda_{a,o}} = 0$, then $M = \theta(r - c_1 - c_2)$ and $N = (1 - \theta)(r - c_2)$ can be obtained. Therefore, the optimal capacity of each stage can be obtained by $\frac{\partial L}{\partial \mu_{a,1}} = 0$ and $\frac{\partial L}{\partial \mu_{a,2}} = 0$ as follows:

By solving the above formula, the optimal capacity at each stage can be obtained as:

$$\begin{aligned} \mu_{a,1} &= \theta\lambda_{a,s} + \sqrt{\frac{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}{c_1}}, \\ \mu_{a,2} &= \theta\lambda_{a,s} + (1 - \theta)\lambda_o + \sqrt{\frac{\theta\beta(r - c_1 - c_2) + (1 - \theta)\epsilon\beta(r - c_2)}{c_2}}, \end{aligned}$$

Then plug the above optimal capabilities into $\frac{\partial L}{\partial M} = 0$ and $\frac{\partial L}{\partial N} = 0$ to obtain the following customer shopping rate function:

$$\begin{aligned} \lambda_{a,s} &= \alpha - (\beta + \delta_s)\sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}} - \beta\sqrt{\frac{c_2}{\theta\beta(r - c_1 - c_2) + (1 - \theta)\epsilon\beta(r - c_2)}} \\ &\quad + \delta_s t > 0, \\ \lambda_{a,o} &= \alpha - \epsilon\beta\sqrt{\frac{c_2}{\theta\beta(r - c_1 - c_2) + (1 - \theta)\epsilon\beta(r - c_2)}} + \delta_o\sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}} \\ &\quad - (\epsilon\beta + \delta_o)t > 0 \end{aligned}$$

The above effective shopping rates $\lambda_{a,s}$ and $\lambda_{a,o}$ must be positive in actual operation. Therefore, according to $\lambda_{a,s} > 0$ and $\lambda_{a,o} > 0$, the effective range of the base shopping rate α with practical significance can be obtained as follows:

$$\begin{cases} \alpha > (\beta + \delta_s)\sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}} + \beta\sqrt{\frac{c_2}{\theta\beta(r - c_1 - c_2) + (1 - \theta)\epsilon\beta(r - c_2)}} - \delta_s t, \\ \alpha > \epsilon\beta\sqrt{\frac{c_2}{\theta\beta(r - c_1 - c_2) + (1 - \theta)\epsilon\beta(r - c_2)}} - \delta_o\sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}} + (\epsilon\beta + \delta_o)t. \end{cases}$$

Therefore, when the base shopping rate α meets the above formula, the effective shopping rate of customers is positive and has significance, i.e., $\lambda_{a,s} > 0$ and $\lambda_{a,o} > 0$. However, if α does not meet the above formula, it does not meet the practical significance

and is not included in the research scope of this paper. We also make a corresponding explanation at the end of Proposition 2. \square

Proof of Lemma 1. We know $\mu_{n,1}^{safe} = \sqrt{\frac{\theta(\beta+\delta_s)(r-c_1-c_2)}{c_1}}$ and $\mu_{n,2}^{safe} = \sqrt{\frac{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}{c_2}}$. Thus the derivative of $\mu_{n,1}^{safe}$ with respect to δ_s is $\frac{\partial\mu_{n,1}^{safe}}{\partial\delta_s} = \frac{1}{2}\left(\frac{\theta(\beta+\delta_s)(r-c_1-c_2)}{c_1}\right)^{-\frac{1}{2}}\frac{\theta(r-c_1-c_2)}{c_1} > 0$, and the derivative of $\mu_{n,1}^{safe}$ with respect to δ_o is $\frac{\partial\mu_{n,1}^{safe}}{\partial\delta_o} = 0$. Meanwhile, the derivative of $\mu_{n,2}^{safe}$ with respect to δ_s is $\frac{\partial\mu_{n,2}^{safe}}{\partial\delta_s} = \frac{1}{2}\left(\frac{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}{c_2}\right)^{-\frac{1}{2}}\frac{\theta(r-c_1-c_2)}{c_1} > 0$, and the derivative of $\mu_{n,2}^{safe}$ with respect to δ_o is $\frac{\partial\mu_{n,2}^{safe}}{\partial\delta_o} = \frac{1}{2}\left(\frac{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}{c_2}\right)^{-\frac{1}{2}}\frac{(1-\theta)(r-c_2)}{c_1} > 0$. \square

Proof of Lemma 2. (1) It is known that when channel information is unavailable, the shopping rate of offline customers can be expressed as follows: $\lambda_{n,s} = \alpha - (\beta + \delta_s) \sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)}} - (\beta + \delta_s) \sqrt{\frac{c_2}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}} + \delta_s \bar{W}_s > 0$. Using the equation to differentiate δ_s , we obtain $\frac{\partial\lambda_{n,s}}{\partial\delta_s} = -\frac{1}{2}\sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)}} - \sqrt{\frac{c_2}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}} \left(\frac{(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)+\frac{1}{2}\theta(\beta+\delta_s)(r-c_1-c_2)}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}\right) + \bar{W}_s$.

According to the above equation $\frac{\partial\lambda_{n,s}}{\partial\delta_s}$, the first two items must be less than 0, and \bar{W}_s must be greater than 0. Therefore, when \bar{W}_s is small, i.e., $\bar{W}_s < \frac{1}{2}\sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)}} + \sqrt{\frac{c_2}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}} \left(\frac{(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)+\frac{1}{2}\theta(\beta+\delta_s)(r-c_1-c_2)}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}\right)$, $\frac{\partial\lambda_{n,s}}{\partial\delta_s}$ must be negative. When \bar{W}_s is large, i.e., $\bar{W}_s > \frac{1}{2}\sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)}} + \sqrt{\frac{c_2}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}} \left(\frac{(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)+\frac{1}{2}\theta(\beta+\delta_s)(r-c_1-c_2)}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}\right)$, $\frac{\partial\lambda_{n,s}}{\partial\delta_s}$ must be positive.

(2) Similarly, we know that the shopping rate of online customers is $\lambda_{n,o} = \alpha - (\epsilon\beta + \delta_o) \sqrt{\frac{c_2}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}} - (\epsilon\beta + \delta_o)t + \delta_o \bar{W}_o$. Therefore, by using $\lambda_{n,o}$ to solve the first-order derivative of the reference point sensitivity of online customers, the following equation can be obtained:

$$\frac{\partial\lambda_{n,o}}{\partial\delta_s} = -\sqrt{\frac{c_2}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}} \left(\frac{\frac{1}{2}(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)+\theta(\beta+\delta_s)(r-c_1-c_2)}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}\right) - t + \bar{W}_o$$

From the above equation, when \bar{W}_o is relatively small, $\frac{\partial\lambda_{n,o}}{\partial\delta_s}$ must be negative and $\frac{\partial\lambda_{n,o}}{\partial\delta_s}$ is positive when \bar{W}_o is relatively large. \square

The proof of Lemma 3. We know that the profits of omnichannel catering firms can be expressed as follows when channel information is unavailable:

$$\pi_n = \theta(\alpha + \delta_s \bar{W}_s) (r - c_1 - c_2) + (1 - \theta)(\alpha - (\epsilon\beta + \delta_o)t + \delta_o \bar{W}_o)(r - c_2) - 2\sqrt{c_1\theta(\beta + \delta_s)(r - c_1 - c_2)} - 2\sqrt{c_2(\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\epsilon\beta + \delta_o)(r - c_2))}$$

Taking the derivative of δ_s in the above equation, $\frac{\partial\pi_n}{\partial\delta_s}$ can be obtained as follows:

$$\frac{\partial\pi_n}{\partial\delta_s} = \theta(r - c_1 - c_2) \left(\bar{W}_s - \sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2)}} - \sqrt{\frac{c_2}{\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\epsilon\beta + \delta_o)(r - c_2)}} \right)$$

We know that $w_1 = \sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)}}$, $w_2 = \sqrt{\frac{c_2}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}}$. Therefore, the above equation is equivalent to $\frac{\partial\pi_n}{\partial\delta_s} = \theta(r - c_1 - c_2) (\bar{W}_s - w_1 - w_2)$. Then

we obtain $\frac{\partial \pi_n}{\partial \delta_s} < 0$ at $\bar{W}_s < \sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)}} + \sqrt{\frac{c_2}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}}$, and $\frac{\partial \pi_n}{\partial \delta_s} > 0$ at $\bar{W}_s > \sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)}} + \sqrt{\frac{c_2}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}}$.

Then, we take the derivative of δ_o in the above equation, $\frac{\partial \pi_n}{\partial \delta_o}$ can be obtained as follows:

$$\frac{\partial \pi_n}{\partial \delta_o} = (1 - \theta)(r - c_2) \left(\bar{W}_o - t - \sqrt{\frac{c_2}{\theta(\beta+\delta_s)(r-c_1-c_2)+(1-\theta)(\epsilon\beta+\delta_o)(r-c_2)}} \right)$$

Therefore, the above equation is equivalent to $\frac{\partial \pi_n}{\partial \delta_o} = (1 - \theta)(r - c_2)(\bar{W}_o - t - w_2)$.

Then we can find $\frac{\partial \pi_n}{\partial \delta_o} < 0$ at $t + w_2 > \bar{W}_o$, and $\frac{\partial \pi_n}{\partial \delta_o} > 0$ at $t + w_2 < \bar{W}_o$. □

Proof of Lemma 4. We know $\mu_{a,1}^{safe} = \sqrt{\frac{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}{c_1}}$ and $\mu_{a,2}^{safe} = \sqrt{\frac{\theta\beta(r-c_1-c_2)+(1-\theta)\epsilon\beta(r-c_2)}{c_2}}$. Thus the derivative of $\mu_{a,1}^{safe}$ with respect to δ_s is $\frac{\partial \mu_{a,1}^{safe}}{\partial \delta_s} = \frac{1}{2} \left(\frac{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}{c_1} \right)^{-\frac{1}{2}} \frac{\theta(r-c_1-c_2)}{c_1} > 0$, and the derivative of $\mu_{a,1}^{safe}$ with respect to δ_o is $\frac{\partial \mu_{a,1}^{safe}}{\partial \delta_o} = -\frac{1}{2} \left(\frac{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}{c_1} \right)^{-\frac{1}{2}} \frac{\theta(r-c_1-c_2)}{c_1} < 0$. Meanwhile, the derivative of $\mu_{a,2}^{safe}$ with respect to δ_s is $\frac{\partial \mu_{a,2}^{safe}}{\partial \delta_s} = 0$, and the derivative of $\mu_{a,2}^{safe}$ with respect to δ_o is $\frac{\partial \mu_{a,2}^{safe}}{\partial \delta_o} = 0$. □

Proof of Lemma 5. (1) We know that the shopping rate of offline channels when information is available is $\lambda_{a,s} = \alpha - (\beta + \delta_s) \sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}} - \beta \sqrt{\frac{c_2}{\theta\beta(r-c_1-c_2)+(1-\theta)\epsilon\beta(r-c_2)}} + \delta_s t$. Therefore, taking the derivative of δ_s from the equation $\lambda_{a,s}$, we obtain

$$\frac{\partial \lambda_{a,s}}{\partial \delta_s} = \frac{-\sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}} + \frac{1}{2}(\beta + \delta_s) \sqrt{\frac{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}{c_1}}}{c_1\theta(r-c_1-c_2)} + t.$$

The above equation $\frac{\partial \lambda_{a,s}}{\partial \delta_s}$ can be further rewritten as:

$$\begin{aligned} \frac{\partial \lambda_{a,s}}{\partial \delta_s} &= \sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}} \left(\frac{1}{2}(\beta + \delta_s) \frac{\theta(r-c_1-c_2)}{(\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2))} \right. \\ &\quad \left. - 1 \right) + t = \sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}} \left(\frac{(1-\theta)\delta_o(r-c_2) - \frac{1}{2}(\beta+\delta_s)(r-c_1-c_2)}{(\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2))} \right) + t \end{aligned}$$

In order for the function to make sense, the term $\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)$ in the square root must be positive. So, if $(1 - \theta)\delta_o(r - c_2) - \frac{1}{2}\theta(\beta + \delta_s)(r - c_1 - c_2) > 0$, i.e., $\frac{2\delta_o}{\beta + \delta_s} \geq \frac{\theta(r - c_1 - c_2)}{(1 - \theta)(r - c_2)}$, then $\frac{\partial \lambda_{a,s}}{\partial \delta_s} > 0$, and when $\frac{2\delta_o}{\beta + \delta_s} < \frac{\theta(r - c_1 - c_2)}{(1 - \theta)(r - c_2)}$, and when t is

less than threshold \bar{t}_1 , we obtain $\frac{\partial \lambda_{a,s}}{\partial \delta_s} < 0$, where $\bar{t}_1 = -\sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}} \left(\frac{(1-\theta)\delta_o(r-c_2) - \frac{1}{2}\theta(\beta+\delta_s)(r-c_1-c_2)}{(\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2))} \right)$. This threshold will change as the δ_s and δ_o . Then

$\frac{\partial \lambda_{a,s}}{\partial \delta_o} < 0$ will always be satisfied.

(2) We know that the shopping rate of online channels when information is available is $\lambda_{a,o} = \alpha - \epsilon\beta \sqrt{\frac{c_2}{\theta\beta(r-c_1-c_2)+(1-\theta)\epsilon\beta(r-c_2)}} + \delta_o \sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}} - (\epsilon\beta + \delta_o)t$. Therefore, taking the derivative of δ_o from the equation $\lambda_{a,o}$, we obtain the following.

$$\begin{aligned} \frac{\partial \lambda_{a,o}}{\partial \delta_o} &= \sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}} + \\ &\quad \frac{1}{2} \sqrt{\frac{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}{c_1}} \frac{c_1(1-\theta)\delta_o(r-c_2)}{(\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2))^2} - t \end{aligned}$$

The above equation can be further rewritten as:

$$\begin{aligned} \frac{\partial \lambda_{a,o}}{\partial \delta_o} &= \sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}} \left(\frac{\frac{1}{2}(1-\theta)\delta_o(r-c_2)}{(\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2))} + 1 \right) - t \\ &= \sqrt{\frac{c_1}{\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2)}} \left(\frac{\theta(\beta+\delta_s)(r-c_1-c_2) - \frac{1}{2}(1-\theta)\delta_o(r-c_2)}{(\theta(\beta+\delta_s)(r-c_1-c_2)-(1-\theta)\delta_o(r-c_2))} \right) - t \end{aligned}$$

In order for the function to make sense, the term $\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)$ in the square root must be positive. So, if $\theta(\beta + \delta_s)(r - c_1 - c_2) - \frac{1}{2}(1 - \theta)\delta_o(r - c_2) < 0$, i.e., $\frac{\delta_o}{2(\beta + \delta_s)} \geq \frac{\theta(r - c_1 - c_2)}{(1 - \theta)(r - c_2)}$, then $\frac{\partial \lambda_{a,o}}{\partial \delta_o} < 0$, and when $\frac{\delta_o}{2(\beta + \delta_s)} < \frac{\theta(r - c_1 - c_2)}{(1 - \theta)(r - c_2)}$, and when t is less than threshold \bar{t}_2 , we obtain $\frac{\partial \lambda_{a,o}}{\partial \delta_o} > 0$, where $\bar{t}_2 = \sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}}$ $\left(\frac{\theta(\beta + \delta_s)(r - c_1 - c_2) - \frac{1}{2}(1 - \theta)\delta_o(r - c_2)}{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}\right)$. This threshold will change as the δ_s and δ_o , and $\frac{\partial \lambda_{a,s}}{\partial \delta_s} < 0$ will always be satisfied. \square

Proof of Lemma 6. It is known that the optimal profit of omnichannel catering firms under the condition that information is available is as follows:

$$\pi_a = \theta(\alpha + \delta_s t) \frac{(r - c_1 - c_2) + (1 - \theta)(\alpha - (\epsilon\beta + \delta_o)t)(r - c_2)}{-2\sqrt{c_1(\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2))} - 2\sqrt{c_2(\theta\beta(r - c_1 - c_2) + (1 - \theta)\epsilon\beta(r - c_2))}}$$

By differentiating δ_s and δ_o in the above equation, $\frac{\partial \pi_a}{\partial \delta_s}$ and $\frac{\partial \pi_a}{\partial \delta_o}$ can be obtained as follows:

$$\frac{\partial \pi_a}{\partial \delta_s} = \theta(r - c_1 - c_2) \left(t - \sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}} \right)$$

$$\frac{\partial \pi_a}{\partial \delta_o} = (1 - \theta)(r - c_2) \left(-t + \sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}} \right)$$

We know that $w_1 = \sqrt{\frac{c_1}{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}}$. Thus, we can obtain $\frac{\partial \pi_a}{\partial \delta_s} < 0$ and 0 when $w_1 > t$ and $\frac{\partial \pi_a}{\partial \delta_s} > 0$, and $\frac{\partial \pi_a}{\partial \delta_o} < 0$ when $w_1 < t$. \square

Proof of Lemma 7. We know that the safety capacities when information is unavailable are $\mu_{n,1}^{safe} = \sqrt{\frac{\theta(\beta + \delta_s)(r - c_1 - c_2)}{c_1}}$ and $\mu_{n,2}^{safe} = \sqrt{\frac{\theta(\beta + \delta_s)(r - c_1 - c_2) + (1 - \theta)(\epsilon\beta + \delta_o)(r - c_2)}{c_2}}$. When information is available, the safety capacities are $\mu_{a,1}^{safe} = \sqrt{\frac{\theta(\beta + \delta_s)(r - c_1 - c_2) - (1 - \theta)\delta_o(r - c_2)}{c_1}}$ and $\mu_{a,2}^{safe} = \sqrt{\frac{\theta\beta(r - c_1 - c_2) + (1 - \theta)\epsilon\beta(r - c_2)}{c_2}}$. Since $(1 - \theta)\delta_o(r - c_2) > 0$, there must be $\mu_{a,1}^{safe} < \mu_{n,1}^{safe}$. However, for $\mu_{n,2}^{safe}$ and $\mu_{a,2}^{safe}$, since δ_o and δ_s are both greater than 0, there must be $\mu_{a,2}^{safe} < \mu_{n,2}^{safe}$. \square

References

1. Anand, K.S.; Paç, M.F.; Veeraraghavan, S. Quality–Speed Conundrum: Trade-offs in Customer-Intensive Services. *Manag. Sci.* **2011**, *57*, 40–56. [CrossRef]
2. Zhan, W.; Jiang, M.; Li, C. The service strategy of customer-intensive services under advertising effects. *Kybernetes* **2020**, *50*, 2453–2470. [CrossRef]
3. Yuan, X.; Dai, T.; Chen, L.G.; Gavirneni, S. Co-Opetition in Service Clusters with Waiting-Area Entertainment. *Manuf. Serv. Oper. Manag.* **2021**, *23*, 106–122. [CrossRef]
4. Gao, F.; Su, X. Omnichannel Service Operations with Online and Offline Self-Order Technologies. *Manag. Sci.* **2018**, *64*, 3595–3608. [CrossRef]
5. Zhan, W.; Jiang, M.; Wang, X. Optimal Capacity Decision-Making of Omnichannel Catering Merchants Considering the Service Environment Based on Queuing Theory. *Systems* **2022**, *10*, 144. [CrossRef]
6. Bell, D.R.; Gallino, S.; Moreno, A. Offline showrooms in omni-channel retail: Demand and operational benefits. *Manag. Sci.* **2018**, *64*, 1629–1651. [CrossRef]
7. Kopalle, P.K.; Winer, R.S. A dynamic model of reference price and expected quality. *Mark. Lett.* **1996**, *7*, 41–52. [CrossRef]
8. Popescu, I.; Wu, Y. Dynamic Pricing Strategies with Reference Effects. *Oper. Res.* **2007**, *55*, 413–429. [CrossRef]
9. Nasiry, J.; Popescu, I. Dynamic Pricing with Loss-Averse Consumers and Peak-End Anchoring. *Oper. Res.* **2011**, *59*, 1361–1368. [CrossRef]

10. Roels, G.; Su, X. Optimal Design of Social Comparison Effects: Setting Reference Groups and Reference Points. *Manag. Sci.* **2014**, *60*, 606–627. [[CrossRef](#)]
11. Yang, L.; Guo, P.; Wang, Y. Service Pricing with Loss-Averse Customers. *Oper. Res.* **2018**, *66*, 761–777. [[CrossRef](#)]
12. Wang, Z.G.; Yang, Y.X.; Liu, H.; Lin, M.Y. Analysis of Consumers' cognition level, reference behavior and benefit degree of agricultural product certification labeling: Based on the questionnaire survey of 20 provinces and autonomous regions in China. *J. Agric. Econ. Manag.* **2013**, *59*, 38–44.
13. Fehr, E.; Schmidt, K.M. A Theory of Fairness, Competition, and Cooperation. *Q. J. Econ.* **1999**, *114*, 817–868. [[CrossRef](#)]
14. Kahneman, D.; Tversky, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* **1979**, *47*, 263. [[CrossRef](#)]
15. Blettner, D.P.; Gollisch, S. The role of reference points and organizational identity in strategic adaptation to performance feedback. *Manag. Res. Rev.* **2022**, *45*, 1205–1228. [[CrossRef](#)]
16. Wang, Y.; Liu, P.; Yao, Y. BMW-TOPSIS: A generalized TOPSIS model based on three-way decision. *Inf. Sci.* **2022**, *607*, 799–818. [[CrossRef](#)]
17. Qiu, R.; Yu, Y.; Sun, M. Joint pricing and stocking decisions for a newsvendor problem with loss aversion and reference point effect. *Manag. Decis. Econ.* **2020**, *42*, 275–288. [[CrossRef](#)]
18. Long, X.; Nasiry, J. Prospect Theory Explains Newsvendor Behavior: The Role of Reference Points. *Manag. Sci.* **2015**, *61*, 3009–3012. [[CrossRef](#)]
19. Kirshner, S.N.; Ovchinnikov, A. Heterogeneity of reference points in the competitive newsvendor problem. *Manuf. Serv. Oper. Manag.* **2018**, *21*, 571–581. [[CrossRef](#)]
20. Zong, J.C. Comparison trap under stereotype: Experimental research on product bundling strategy. *Nankai Manag. Rev.* **2018**, *21*, 210–218.
21. Verhoef, P.C.; Kannan, P.; Inman, J.J. From Multi-Channel Retailing to Omni-Channel Retailing. *J. Retail.* **2015**, *91*, 174–181. [[CrossRef](#)]
22. Narang, U.; Shankar, V. Mobile App Introduction and Online and Offline Purchases and Product Returns. *Mark. Sci.* **2019**, *38*, 756–772. [[CrossRef](#)]
23. Gallino, S.; Moreno, A. Integration of Online and Offline Channels in Retail: The Impact of Sharing Reliable Inventory Availability Information. *Manag. Sci.* **2014**, *60*, 1434–1451. [[CrossRef](#)]
24. Bell, D.; Gallino, S.; Moreno, A. Showrooms and Information Provision in Omni-channel Retail. *Prod. Oper. Manag.* **2014**, *24*, 360–362. [[CrossRef](#)]
25. Chopra, S. How omnichannel can be the future of retailing. *Decision* **2016**, *43*, 135–144. [[CrossRef](#)]
26. Yan, L.; Guo, B.; Chao, C.; He, D.; Yu, Z.; Zhang, D. Foodnet: Toward an optimized food delivery network based on spatial crowdsourcing. *IEEE Trans. Mob. Comput.* **2018**, *6*, 1288–1301.
27. Pei, Y.; Li, D.; Xue, W. The evaluation of customer experience using BP neural network-taking catering O2O takeout. *Concurr. Comput. Pract. Exp.* **2019**, *32*, e5515. [[CrossRef](#)]
28. Zhao, D.Z.; Yang, J. Research on Decision-making Model of Estimated Delivery Time in O2O Food Delivery. *Ind. Eng. Neering Management.* **2018**, *23*, 8–15.
29. Chen, M.; Hu, M.; Wang, J. Food Delivery Service and Restaurant: Friend or Foe? *Manag. Sci.* **2022**. [[CrossRef](#)]
30. Bensoussan, A.; Guo, P. Technical Note—Managing Nonperishable Inventories with Learning About Demand Arrival Rate Through Stockout Times. *Oper. Res.* **2015**, *63*, 602–609. [[CrossRef](#)]
31. Gao, F.; Su, X. Online and Offline Information for Omnichannel Retailing. *Manuf. Serv. Oper. Manag.* **2017**, *19*, 84–98. [[CrossRef](#)]
32. Li, W.; Jiang, M.; Zhan, W. Research on the Pricing Decisions of a Video Platform Based on Interaction. *Systems* **2022**, *10*, 192. [[CrossRef](#)]
33. Kostecki, M. Waiting lines as a marketing issue. *Eur. Manag. J.* **1996**, *14*, 295–303. [[CrossRef](#)]
34. Bassamboo, A.; Randhawa, R.S.; Zeevi, A. Capacity Sizing Under Parameter Uncertainty: Safety Staffing Principles Revisited. *Manag. Sci.* **2010**, *56*, 1668–1686. [[CrossRef](#)]
35. Guo, W.; Chen, X.-R.; Liu, H.-C. Decision-Making under Uncertainty: How Easterners and Westerners Think Differently. *Behav. Sci.* **2022**, *12*, 92. [[CrossRef](#)] [[PubMed](#)]