RAIRO-Oper. Res. 58 (2024) 4277–4305
https://doi.org/10.1051/ro/2024149

RAIRO Operations Research
www.rairo-ro.org

TAKE-AWAY AND SIT-DOWN SERVICE OPERATIONS UNDER INEQUITY AVERSION

Sijia Zeng[1], Wei Hu[2] and Xiaochen Sun[3,*]

Abstract. Profit contributed by take-away service has become an increasingly essential element of the restaurant operating revenue. Since take-away service usually relies on the third-party platform and there are many differences between the cost of take-away service and that of sit-down service. We focus on the restaurant which provides both sit-down service and take-away service, the service system is modeled as a two-stage tandem queueing system. We study the restaurant's optimal capacity level for each stage. Besides, as there exist price/waiting time difference between the two services, inequity aversion is also investigated in our model. We study symmetrical and asymmetrical inequity aversion. We find that the optimal service capacity consists of two parts, base capacity and safety capacity. And the loss resulted from waiting time lag is equal to the waste of resources caused by fluctuations in arrival rate. Further, when customers really long for the restaurant, even if high price will lead to severe inequity aversion, restaurant can always earn more by raising the price in service channel with high revenue. While when customers are indifferent of the restaurant and the others, the price gap is meant to result in revenue decrease. In addition, reduction in customers' susceptibility can help to enhance operation profit in general. Market environment plays a decisive role in choosing the optimal service level.

Mathematics Subject Classification. 90B22, 90B50.

Received May 21, 2023. Accepted July 18, 2024.

1. Introduction

Statistics from National Bureau of Statistics of China show that in 2018, China's catering industry revenue growth rate exceeds that of retail sales, catering delivery and take-away service are growing at a particularly fast rate. Take-away service helps consumers to save the time for finding restaurants, ordering online and waiting for meals. Because a large quantity of people choose to dine out, companies and restaurants are developing technologies to make it easier to serve an ever-growing customer-base. Those technologies include phone applications, online delivery services, and online customer review sites. Meanwhile, restaurants can also benefit from it, since the online platform brings additional revenue, improves reputability and does not require an increase in the size of the counter service [24]. The monopoly of the network take-out platform exploits the profits of the

Keywords. Take-away service, sit-down service, capacity management, inequity aversion.

[1] Department of Mathematical Sciences of Tsinghua University, Beijing 100084, P.R. China.
[2] School of Mathematical Sciences and LPMC, Nankai University, Tianjin 300071, P.R. China.
[3] College of Management and Economics, Tianjin University, Tianjin 300072, P.R. China.
*Corresponding author: xcsun@tju.edu.cn

merchants by charging commissions and shipping fee. In the current market, some chain restaurants, such as MacDonald, KFC and Pizza Hut, can develop delivery system independently relying on their sufficient footfall and capital. While most small and medium-sized restaurants do not have such capabilities. They may still pay a high commission for the third platform, or deliver by themselves in a small range through other platforms such as WeChat.

We consider a restaurant deals with two types of customers: (i) take-away customers, who place orders *via* the online self-order channel, wait for the food to be prepared and delivered; (ii) sit-down customers, who physically come to the store to place an order and wait for the food to be prepared. The sales price of different channels is bound to be different because of sales cost, price difference is an important part of revenue management.

The inequity aversion originated from behavioral economics has been confirmed by many studies and is widely used in marketing pricing, behavioral decision-making, etc. [9]. Inequity aversion refers to the preference for fairness and the resistance to significant inequality, and is often discussed in the economic and social literatures. Customers usually compare the price of a commodity with its historical price or the price of other alternatives, and will be dissatisfied with the possible price difference, which will greatly affect their evaluation of the business [16]. Inequity aversion affects revenue through two possible ways: one is that when consumers have low elasticity of demand for products and it is difficult to find alternatives, the price difference between channels drives consumers to switch from high-cost channel to low-cost channel; the other is that when consumers have high elasticity of demand for products, they are more likely to choose the alternatives. We study symmetrical inequity aversion and asymmetrical inequity aversion according to consumers' elasticity of demand. There is no doubt that comparisons of prices will affect customers' consumption decisions.

We develop a stylized model, there is a restaurant serving wait-sensitive customers. The service system is modeled as a two-stage tandem queueing network, where customers first place an order through the register and then wait for their order to be prepared in the kitchen. The first stage is the process that consumers order meals at the counter and the second stage is the process that kitchen caters to two types of consumers. The two stages both can be regarded as $M/M/1$ queueing systems. Time-sensitive consumers' utility will change due to prices and waiting time, and their utility is closely related to customer arrival rate per unit time. In our model, the operator selects the capacity level of each stage to maximize the revenue, and the behavior of the consumer and the third-party platform is viewed as an exogenous variable. However, in the practical market, consumers who get take-away service do face price/waiting time discrimination, and the price/waiting time difference between the two channels have a certain impact on consumer behavior. We introduce a conception from behavioral economics, inequity aversion, and explain the result economically through the rule of thumb for capacity planning. Based on a basic two stage capacity optimization problem, we further discuss two-stage optimal capacity considering symmetric and asymmetric inequity aversion may caused by price or waiting time. Moreover, based on the basic model, we investigate two additional discussion: (1) the restaurant has to decide optimal price and capacity levels for the take-away and sit-down channel; (2) the self-delivery model, namely, the restaurant has to decide the optimal delivery time and capacity levels.

In this paper we focus on the following research questions:

(1) As a market price receiver, how the restaurant choose the optimal capacity level for the counter in the first stage and the kitchen in the second stage?
(2) Facing with symmetric and asymmetric inequity aversion caused by price or waiting time, how the restaurant choose the optimal capacity level for two stages?
(3) How exogenous factors determined by the market and consumers affect the optimal decision and income of restaurant?

At present, the researches on the take-away service mainly focus on multi-channel supply chain management under inequity aversion [13, 29] and the O2O business model of take-away platform [34]. The catering industry is different from the general retail industry because of its timeliness, so the concept of time cost is introduced in the model, and the delivery process is also characterized in the model. In the above mentioned literatures, there are few models calculating total income based on the queueing system, we simulate the operation of restaurant more practically using queueing theory. Our work mainly contributes to examining the impact of inequity aversion on restaurants' capacity management from the perspective of behavioral economics, analyzing the impact mechanism and effects of exogenous variables on optimal capacity levels. We also present reasonable economic explanation for each part, such as basic capacity, safety capacity and fluctuation losses. Besides, we explore how the exogenous variables such as waiting time sensitivity affect the total income in different situations.

We present a literature review in Section 2. We describe the model and provide a benchmark model in Section 3. Both the symmetry and asymmetry inequity aversion caused by price/waiting time are studied in Section 4, we also derive managerial insights in this section. In Section 5, we conduct additional discussions. Concluding remarks and further research are provided in Section 6.

2. Literature review

Our study considers a restaurant provides both sit-down service and take-away service to the inequity-aversion customers. There are two main streams of literature most closely related to our work: (1) capacity management in a queueing system; and (2) inequity aversion in pricing management.

The first stream of research is related to capacity management in a queueing system. There is a large body of literature studying capacity management in a queueing system where customers are sensitive to waiting. Hall *et al.* [18] believe that in the queuing model, the completion service time is random, and the price is used to adjust the arrival rate of the queuing system. Bassamboo *et al.* [6] study resource allocation problem in a single-queue multi-service counter system under uncertain parameters. Assuming that demand is random and sensitive to both service and price, using queuing theory, Hwang *et al.* [24] construct an optimization model to solve the problem of demand and capacity management. Based on Markov process model, Ackere *et al.* [2] considering their estimated waiting time, customers decide whether to stay, and operators determine capacity level based on their estimated demand. Introducing the concept of time cost and assuming that customers are sensitive to price and waiting time, Haviv and Randhawa [20] use a $M/M/1$ queue model to simulate service process. When queueing model is extended to $GI/GI/1$, Lee and Ward [27] investigate how to maximize the expected income by jointly controlling price and capacity in a queue with high customer arrival rates. Kurz [26] introduces time limit requirements and timeout penalties, discuss the issue of minimizing the cost of producers in a periodic production networks. Aiming at investigating waiting behavior under e-commerce environment, Hum *et al.* [23] define the concept of supply chain responsiveness as the probability of satisfying customer demands and study how to apply the queuing network model to optimize supply chain responsiveness. In the make-to-order stochastic systems with an endogenous demand built by Albana *et al.* [4], they consider the lead time quotation problem and model the unit operating cost as a convex decreasing function of the quoted lead time. Considering a two-stage tandem queueing model, Gao and Su [15, 16] study effect of the omnichannel service operation with self-order technology on customer demands, employment level and total revenue. Zhan *et al.* [33] focus on analyzing the capacity management of omnichannel catering merchants at the ordering, production and delivery stages. In all the above mentioned papers, researchers do not involve the impact of inequity aversion. We contributed to the allocation theory of capacity management by examining the impact of inequity aversion from the perspective of behavioral operation research.

The second relevant literature concerns research on inequity aversion problem. Inequity aversion, also known as fairness preference, which deals with social preferences mainly appears in behavioral economics and has a lot of application in supply chain. We focus on two generic inequity aversion models originally developed by Fehr and Schmidt [14] and Bolton and Ockenfels [8], and discuss their application to capacity management. Chen and Cui [11] prove that in the price competition model that incorporates inequity aversion, equilibrium will achieve for the same type of products. And the unfair pricing has a significant negative impact on the effectiveness and evaluation of customers, which is verified thoroughly by Heo *et al.* [21]. Although the sales price of different channels is bound to be different because of sales cost, controlling and coordinating sales, price difference is still an important part of revenue management. In Guo's work [17], consumer perception of price fairness is considered to be uncertain and susceptible, and under this assumption he derives a superior sales strategy. We study symmetrical inequity aversion and asymmetrical inequity aversion according to consumers' elasticity of demand.

Besides, there are some works related to two-stage tandem queueing system. Medhi [28] introduces queues in series as a model of a queuing system, and proves that in steady state each stage behaves independently of the other. Ahn *et al.* [3] consider the optimal control of two parallel servers in a two-stage tandem queueing system. Papachristos and Padelis [30] study server allocation strategies in a queueing system consisting of two stations in tandem, and determine structural properties of the optimal strategies. In our model, we emphasize the influence of capacity allocation on arrival rate and revenue.

Above all, we solve the two stage capacity optimization problem, and expand the basic model by examing symmetric and asymmetric inequity aversion may caused by price or waiting time, which enriches the capacity management theory in omnichannel.

3. Basic model

There is a restaurant serving food to customers. The service system is modeled as a two-stage tandem queueing network: at stage 1, customers place their orders through the front-end cashiers/waiters, at stage 2, the order is transmitted to the back-end kitchen area where food is prepared. The two stages both can be regarded as M/M/1 queueing systems. The firm decides the service rate (or capacity) at each stage, denoted by μ_1 and μ_2.

In our model, the restaurant deals with two types of customers: (i) take-away customers, who place orders *via* the online self-order channel, wait for the food to be prepared and delivered; (ii) sit-down customers, who physically come to the store to place an order and wait for the food to be prepared. Suppose the fraction of the take-away customers in the market is $\theta \in (0,1)$, and the remaining $1-\theta$ are sit-down customers. Restaurant service includes both ordering and food preparation stages; both require customers to wait. Take-away customers can instantaneously place orders through their own digital devices and do not need to wait in line for their turn to order the food. Thus their wait time w_1 at stage 1 is reduced to 0. Customers still have waiting time w_2 for food to be prepared in kitchen and waiting time t for delivering, their total waiting time thus is w_2+t. For sit-down customers, they wait in line at stage 1 to place their order and their waiting time is w_1, and they also have waiting time for food to be prepared in the kitchen which is w_2. As a result, their total waiting time in the store is the sum of the waiting time at both stages is w_1+w_2. Assuming arrival time interval of customers and preparation time interval of restaurant are subject to the Poisson distribution and they form a $M/M/1$ queueing system. Given the service capability μ and the arrival rate λ in a $M/M/1$ queue, we know that average waiting time is $\frac{1}{\mu-\lambda}$ in equilibrium.

Consumers come to the restaurant randomly and $1-\theta$ of them choose sit-down service while others order through electronic equipments and wait for food to be delivered. They will pay for the deal only when they get utility larger than costs. The total cost of customers consists of price cost and time cost. The price cost p is the expense for service and time cost $T\beta$, which T is total waiting time. Since consumers are sensitive to time and many expect their orders to be delivered quickly, we assume β is the waiting time sensitivity. Consumers' utility is a random variable V, whose cumulative distribution function is recorded as $F(x)$, and let $\overline{F}(x)=1-F(x)$. Customers have the incentive to consume if and only if their utility is higher than cost, that is to say, a consumer's probability of consuming is $\overline{F}(p+T\beta)$. For the sake of convenience, we assume that the utility V is subject to uniform distribution $U(0,1)$, so $\overline{F}(p+T\beta)=1-(p+\beta T)$ for $p+\beta T<1$, $\overline{F}(p+T\beta)=0$ for $p+\beta T\geq 1$. It means that restaurant must make a price less than 1, otherwise no consumers will come. Let λ represents total arrival rate, then expected effective demand rate is $\lambda\overline{F}(p+T\beta)$. We use the word "effective" here since not all consumers who come to the restaurant will join the queue. Customers learn the price and queue length, which decide their cost, and have incentive to consume if and only if their utility is higher than cost. According to the conclusion of Haviv and Randhawa [20], we assume that the price p_s, p_t are fixed, which are the unit price for unit service of sit-down and take-away consumers respectively. Thus, the arrival rate is $(1-\theta)\lambda\bar{F}(p_s+\beta_sT_s)$ for the sit-down customers, and which is $\theta\lambda\bar{F}(p_t+c_d+\beta_tT_t)$ for take-away customers. In

TABLE 1. Notations.

β_s, β_t [\$/time]	Waiting time sensitivity of sit-down consumer and take-away consumer
λ_s, λ_t [consumers/time]	Effective arrival rate of sit-down consumer and take-away consumer
λ [consumers/time]	The total arrival rate
p_s, p_t [\$]	Price for every unit of sit-down service and take-away service
r_s, r_t [\$]	Income of every unit of sit-down service and take-away service
θ	Proportion of customers who choose to take out
T_s, T_t [time]	Total waiting time of sit-down consumer and take-away consumer
c_1, c_2 [\$]	Cost of every unit of service at stage 1 and 2
c_d [\$]	Shipping fee
δ	Percentage of service fee charged by the third-party platform
μ_1, μ_2 [consumers/time]	Capacity level at stage 1 and 2 (decision variables)
π [\$]	Total operating income

expressions above, β_s and β_t are waiting time sensitivity, c_d is shipping cost paid by the take-away customers for food to be delivered. Because λ is a constant, we standardize the price p_s, p_t and rewrite the arrival rate of sit-down customers and take-away customers as $(1-\theta)[\lambda-(p_s+\beta_s T_s)]$ and $\theta[\lambda-(p_t+c_d+\beta_t T_t)]$ when $\lambda > p_s+\beta_s T_s, \lambda > p_t+c_d+\beta_t T_t$. Based on this parameter adjustment, the quantities p and βT have same dimensions as λ. In the following, parameters p, c_d and β not only measure price of service and time, but also affect arrival rate directly. Their numerical value show relative price, while dimensions represent arrival rate.

Note that because take-away customers skip the ordering stage in the store, the demand at stage 1 includes only sit-down customers. In addition, since restaurants typically do not distinguish take-away and sit-down orders in the kitchen at stage 2. In the counter, we have a queue with service rate μ_1 and arrival rate $(1-\theta)\lambda_s$. In the kitchen, we have a queue with service rate μ_2 and arrival rate $(1-\theta)\lambda_s+\theta\lambda_t$. The waiting time for two kinds of consumers is partly different. The subscript "t" refers to the take-away consumers. T_t is the total waiting time for take-away consumers, which consists of waiting time for kitchen and delivering. T_s is the total waiting time for sit-down consumers, which made up of waiting time for counter and kitchen. λ_s and λ_t are the customers' arrival rate of sit-down and take-away channels. Therefore, according to our model formulation, we have the wait time for take-away and sit-down customers are $T_s=\frac{1}{\mu_1-(1-\theta)\lambda_s}+\frac{1}{\mu_2-(1-\theta)\lambda_s-\theta\lambda_t}$ and $T_t=\frac{1}{\mu_2-(1-\theta)\lambda_s-\theta\lambda_t}+t$ in equilibrium. The delivery is offered by the third platform, the time for delivering t is supposed to be fixed. Meanwhile, the third platform charges a percentage δ of every delivering.

All necessary notations and their units used in this paper are as follows (Tab. 1).

We consider a long-term two-stage restaurant model: the restaurant provides sit-down and take-away service and its service system consists of a front-end counter for ordering and a kitchen for preparation. The revenue of the restaurant is equal to the income brought by the two channels $p_s(1-\theta)\lambda_s+p_t\theta(1-\delta)\lambda_t$ subtracting the total service cost $c_1\mu_1+c_2\mu_2$, where c_1 and c_2 are cost of every unit of service at stage 1 and 2. Given the demand function, the firm chooses the capacity level at each stage, *i.e.* μ_1 and μ_2, to maximize the profit rate. We use superscript $\cdot^b$ (for "basic") to denote the basic model.

$$
\begin{aligned}
\max_{0\le(1-\theta)\lambda_s^b\le\mu_1,\,0\le(1-\theta)\lambda_s^b+\theta\lambda_t^b\le\mu_2} \quad & \left\{p_s(1-\theta)\lambda_s^b+p_t\theta(1-\delta)\lambda_t^b-c_1\mu_1-c_2\mu_2\right\}\\
\text{s.t. } \lambda_s^b &= \left[\lambda-\beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s^b}+\frac{1}{\mu_2-(1-\theta)\lambda_s^b-\theta\lambda_t^b}\right)-p_s\right]^+, \\
\lambda_t^b &= \left[\lambda-\beta_t\left(\frac{1}{\mu_2-(1-\theta)\lambda_s^b-\theta\lambda_t^b}+t\right)-p_t-c_d\right]^+.
\end{aligned} \tag{1}
$$

Considering that in the long run, restaurants will cancel services with little footfall, we add constraints that $\lambda_s, \lambda_t > 0$, *i.e.*

$$\begin{aligned}\lambda &> \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s^b}+\frac{1}{\mu_2-(1-\theta)\lambda_s^b-\theta\lambda_t^b}\right)+p_s,\\ \lambda &> \beta_t\left(\frac{1}{\mu_2-(1-\theta)\lambda_s^b-\theta\lambda_t^b}+t\right)+p_t+c_d.\end{aligned} \tag{2}$$

Let $r_s = (1-\theta)(p_s - c_1 - c_2)$ and $r_t = \theta[p_t(1-\delta) - c_2]$, which are the unit net profit brought by sit-down service and take-away service,then we get the optimal solution in the following proposition.

Proposition 1. *There is a $\overline{\lambda}$ such that when $\lambda > \overline{\lambda}$, there exists a unique optimal solution:*

$$\mu_1^b = (1-\theta)\lambda_s^b + \sqrt{\frac{\beta_s r_s}{c_1}}, \tag{3}$$

$$\mu_2^b = (1-\theta)\lambda_s^b + \theta\lambda_t^b + \sqrt{\frac{\beta_s r_s + \beta_t r_t}{c_2}}, \tag{4}$$

where,

$$\lambda_s^b = \lambda - \beta_s\left(\sqrt{\frac{c_1}{\beta_s r_s}} + \sqrt{\frac{c_2}{\beta_s r_s + \beta_t r_t}}\right) - p_s, \tag{5}$$

$$\lambda_t^b = \lambda - \beta_t\left(\sqrt{\frac{c_2}{\beta_s r_s + \beta_t r_t}} + t\right) - p_t - c_d. \tag{6}$$

At this time, the maximum profit is:

$$\pi^b = r_s\lambda_s^b + r_t\lambda_t^b - \sqrt{c_1\beta_s r_s} - \sqrt{c_2(\beta_s r_s + \beta_t r_t)}. \tag{7}$$

All proofs are shown in appendix, please refer to the proof for the detail of $\overline{\lambda}$.

According to the rule of thumb for capacity planning mentioned by Bassamboo *et al.* [6], the optimal capacity level of front-end counter μ_1^b is composed of two parts: a "base capacity" to satisfy the mean demand $(1-\theta)\lambda_s^b$, which is determined by arrival rate, and a "safety capacity" $\sqrt{\frac{\beta_s r_s}{c_1}}$ to hedges against variability in practical arrivals. Correspondingly, the optimal capacity level of kitchen μ_2^b is made up of base capacity, that is $(1-\theta)\lambda_s^b + \theta\lambda_t^b$, and safety capacity $\sqrt{\frac{\beta_s r_s+\beta_t r_t}{c_2}}$. The total operating revenue π^b can also be viewed as the mean revenue $r_s\lambda_s^b + r_t\lambda_t^b$, which is brought by base capacity, subtracting the loss $\sqrt{c_1\beta_s r_s} + \sqrt{c_2(\beta_s r_s + \beta_t r_t)}$ resulting from arrival variability.

Proposition 2. *The loss of operating income caused by waiting lag is equal to the loss arise from fluctuations in arrival rate.*

The loss caused by the waiting time lag is essentially due to the decrease in the revenue caused by the customer waiting for the service. The loss caused by the fluctuation of the arrival rate is essentially the additional cost incurred by the restaurant waiting for the customer to come. These two kinds of loss are two sides of the same coin. Through the above analysis, we know that the total operating income π_b can also be regarded as the mean revenue $r_s\lambda_s^b + r_t\lambda_t^b$ subtracting operating loss $\sqrt{c_1\beta_s r_s} + \sqrt{c_2(\beta_s r_s + \beta_t r_t)}$ caused by waiting lag.

4. Inequity aversion

According to Walster and Berscheid's theory [31], people have an antipathy towards unfair situations that are unfavorable to them, and the sense of resistance derived from psychological activities directly affects their consumption decisions and evaluation. Referring to the theory proposed by Fehr and Schmidt [14], we introduce the inequity inefficiency factor γ to characterize the change in consumption propensity due to price or waiting time differences, $\frac{1}{\gamma}$ represents the dependence of consumers on the channel, that is, the stableness of demand with regard to price or waiting time changes. The larger γ is, (*i.e.* the smaller $\frac{1}{\gamma}$ is), the larger footfall' s change is, which results from demand elasticity and price or waiting time difference. From the perspective of consumers, γ measures how much they rely on the purchase channel.

In addition, Ho and Su [22] and Chen and Cui [11] believe that there may be two types of conditions when price difference leads to inequity aversion: symmetry inequity aversion and asymmetry inequity aversion. The former will lead to footfall transferring between channels, the latter will lead to a direct reduction in demand, and the type of inequity aversion depends on the substitutability of services. In the following, we will first investigate symmetry inequity aversion, then discuss the asymmetry inequity aversion case.

4.1. Symmetry inequity aversion to price difference

Under the circumstance of symmetry inequity aversion, consumers are less sensitive and have less flexibility in demand for products. Although they resist the price difference between channels, consumers still do not want to give up the goods because the substitutability of the goods is very small. Therefore,the total arrival rate λ remains unchanged, consumers of high-priced channels will move to low-priced channels, the amount of transfer is proportional to the price difference. The practical explanation is: assume that when the consumer finds that the take-away price is higher than the sit-down price, they will prefer to go to the store because of mental resistance to high price of take-away service. The greater the price difference between the two channels, the more likely consumers will choose sit-down service. We use superscript $\cdot^s$ (for symmetry) to denote the case with symmetry inequity aversion, and the amount of transferring is $\gamma(p_t - p_s)$. (if $p_t > p_s$, there are some consumers who were all set to order delivery turning to the entity restaurant; if $p_t < p_s$, the flow direction of consumers is contrary.) We let $\Delta\lambda_s = \frac{\gamma}{1-\theta}(p_t - p_s), \Delta\lambda_t = \frac{\gamma}{\theta}(p_s - p_t)$. $\Delta\lambda_s$ and $\Delta\lambda_t$ are the change amount of arriving rate of sit-down consumer and take-away consumer. Then the optimization problem turns to:

$$\begin{aligned}
&\max_{\mu_1^s,\mu_2^s} \{p_s(1-\theta)\lambda_s^s + p_t\theta(1-\delta)\lambda_t^s - c_1\mu_1^s - c_2\mu_2^s\} \\
&\text{s.t. } \lambda_s^s = \left[\lambda - \beta_s\left(\frac{1}{\mu_1^s - (1-\theta)\lambda_s^s} + \frac{1}{\mu_2^s - (1-\theta)\lambda_s^s - \theta\lambda_t^s}\right) - p_s + \frac{\gamma}{1-\theta}(p_t - p_s)\right]^+, \\
&\qquad \lambda_t^s = \left[\lambda - \beta_t\left(\frac{1}{\mu_2^s - (1-\theta)\lambda_s^s - \theta\lambda_t^s} + t\right) - p_t - c_d + \frac{\gamma}{\theta}(p_s - p_t)\right]^+.
\end{aligned} \tag{8}$$

We get the optimal capacity level μ_1^s and μ_2^s in Proposition 3.

Proposition 3. *When there is symmetry inequity aversion, the optimal capacity levels at two stages become*

$$\mu_1^s = \mu_1^b + \gamma(p_t - p_s), \quad \mu_2^s = \mu_2^b. \tag{9}$$

The optimal revenue is

$$\pi^s = \pi^b + \gamma(p_t - p_s)\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right), \tag{10}$$

where μ_1^b, μ_2^b *and* π^b *are presented as in Proposition 1, respectively.*

Assume that the price of take-away service is higher than that of sit-down service, *i.e.* $p_t - p_s > 0$, on the premise that $\lambda_t^s > 0$, restaurant can avoid loss of revenue if and only if $\theta r_s > (1-\theta) r_t$, *i.e.* $p_s - c_1 > p_t(1-\delta)$. We can consider the result in a more practical way: the transferring of footfall, which arises from inequity aversion and price difference, reduces the demand for take-away service on one hand, and increases the demand for sit-down service on the other hand. The increase in store service revenue can offset the decrease in the revenue from the take-away service if and only if take-away service brings less unit net profit than sit-down service, *vice versa*.

The higher footfall in sit-down channel will necessarily increase capacity level of front-end counter μ_1. If there are more people choosing sit-down service, the positive effect is more significant. The direction of change in capacity level of kitchen is not so explicit since it is impacted by two elements, and we can just declare that it is in the direction of main source of customers. For example, when $\theta > \frac{1}{2}$, take-away channel is the main source and the variation of demand in this channel takes the privilege, which can not be compensated by that in sit-down channel.

In this case, the inequity aversion behaviors, determined by inequity inefficiency factor and price difference, affect footfall immediately and then capacity levels and total revenue. In conclusion, the restaurant should try to increase the unit service revenue brought by the low-priced channel as much as possible, to reduce losses or increase gains, which is in line with the reality.

4.2. Asymmetry inequity aversion to price difference

Under the circumstance of asymmetry inequity aversion, consumers are more sensitive, so the footfall in the high-priced channel will decrease, and footfall in the low-priced channels will remain unchanged. Unlike the condition in symmetry inequity aversion, where the total number of customers stays the same, asymmetry inequity aversion will reduce the total number of customers. Assume that the loss in footfall is still linearly related to price difference. Let $\Delta\lambda_s^a = \frac{\gamma}{1-\theta}(\min(p_s,p_t) - p_s), \Delta\lambda_t^a = \frac{\gamma}{\theta}(\min(p_s,p_t) - p_t)$. $\Delta\lambda_s^a$ and $\Delta\lambda_t^a$ are change amount of arriving rate of sit-down consumer and take-away consumer for the asymmetry inequity aversion case. Thus the optimization problem turns out to be:

$$\begin{aligned} &\max_{\mu_1^a,\mu_2^a} \{p_s(1-\theta)\lambda_s^a + p_t\theta(1-\delta)\lambda_t^a - c_1\mu_1^a - c_2\mu_2^a\} \\ &\text{s.t. } \lambda_s^a = \left[\lambda - \beta_s\left(\frac{1}{\mu_1^a - (1-\theta)\lambda_s^a} + \frac{1}{\mu_2^a - (1-\theta)\lambda_s^a - \theta\lambda_t^a}\right) - p_s + \frac{\gamma}{1-\theta}(\min(p_s,p_t) - p_s)\right]^+, \\ &\qquad \lambda_t^a = \left[\lambda - \beta_t\left(\frac{1}{\mu_2^a - (1-\theta)\lambda_s^a - \theta\lambda_t^a} + t\right) - p_t - c_d + \frac{\gamma}{\theta}(\min(p_s,p_t) - p_t)\right]^+. \end{aligned} \tag{11}$$

Same as the analysis in Section 4.1, we suppose that the price p_s, p_t is fixed, the change of total income π and capacity levels is only relative to the change of arrival rate λ, and we get the following Proposition 4.

Proposition 4. *When there is asymmetry inequity aversion, the optimal capacity levels at two stages become*

$$\mu_1^a = \mu_1^b + \gamma(\min(p_s,p_t) - p_s), \tag{12}$$

$$\mu_2^a = \mu_2^b - \gamma|p_s - p_t|. \tag{13}$$

The optimal revenue becomes

$$\pi^a = \pi^b + r_s\frac{\gamma}{1-\theta}(\min(p_s,p_t) - p_s) + r_t\frac{\gamma}{\theta}(\min(p_s,p_t) - p_t) < \pi^b, \tag{14}$$

where μ_1^b, μ_2^b and π^b are presented as in Proposition 1, respectively.

Let $\Delta\lambda_s^a = \frac{\gamma}{1-\theta}(\min(p_s,p_t) - p_s), \Delta\lambda_t^a = \frac{\gamma}{\theta}(\min(p_s,p_t) - p_t)$, and $\Delta\pi^a = r_s\frac{\gamma}{1-\theta}(\min(p_s,p_t) - p_s) + r_t\frac{\gamma}{\theta}(\min(p_s,p_t) - p_t)$. Due to the loss of footfall, the total operating income and capacity level of kitchen will inevitably decrease. Assume that the price of take-away service is higher than that of sit-down service, *i.e.* $\min(p_s, p_t) = p_s$, on the premise that $\lambda_t^a > 0, \Delta\pi^a = r_t \frac{\gamma}{\theta}(p_s - p_t)$, the loss of total revenue is equal to the loss of income caused by loss of customers in high-priced channel since the inequity aversion result from the price difference only affects the high-priced channel. At this time, $\Delta\lambda^a$ is a linear function of price difference and inequity inefficiency factor γ, and the operator should try to reduce the capacity level of high-priced channel to avoid losses.

The capacity level of kitchen is in the same situation as the total income, because decreasing demand leads to lower capacity level of kitchen. While the capacity level of front-end counter only relies on footfall of sit-down channel and it keeps pace with the change of λ_s.

4.3. Symmetry inequity aversion to time difference

Inequity aversion to price difference is the most common, while inequity aversion to time difference also has great influence on queueing system. Considering the situation that consumers can transfer in two channels conveniently, they will move to channel which requires less waiting time due to inequity aversion. The amount of transfer is proportional to waiting time difference. The practical explanation is: assume that when the consumer finds that they can save time by ordering take-away, they will be dissatisfied to go to the restaurant because of mental resistance to longer wait. The longer the waiting time difference between the two channels is, the more likely consumers are to choose the channel with less wait. We still use superscript $\cdot^s$ (for symmetry) in this section, and employ inequity inefficiency factor η instead of γ to distinguish from Section 4.1. Then we get a problem which is a little different from the basic one:

$$\begin{aligned}
&\max_{\mu_1^s,\mu_2^s} \{p_s(1-\theta)\lambda_s^s + p_t\theta(1-\delta)\lambda_t^s - c_1\mu_1^s - c_2\mu_2^s\} \\
&\text{s.t. } \lambda_s^s = \left[\lambda - \beta_s\left(\frac{1}{\mu_1^s-(1-\theta)\lambda_s^s} + \frac{1}{\mu_2^s-(1-\theta)\lambda_s^s-\theta\lambda_t^s}\right) - p_s + \frac{\eta}{1-\theta}(t - \frac{1}{\mu_1^s-(1-\theta)\lambda_s^s})\right]^+, \qquad (15)\\
&\qquad \lambda_t^s = \left[\lambda - \beta_t\left(\frac{1}{\mu_2^s-(1-\theta)\lambda_s^s-\theta\lambda_t^s} + t\right) - p_t - c_d + \frac{\eta}{\theta}\left(\frac{1}{\mu_1^s-(1-\theta)\lambda_s^s} - t\right)\right]^+.
\end{aligned}$$

The optimal capacity level μ_1 and μ_2, revenue π are presented in Proposition 5.

Proposition 5. *When there is symmetry inequity aversion to waiting time difference, the optimal capacity level at two stages change accordingly.*

$$\begin{aligned}
\Delta\lambda_s^s &= \lambda_s^b - \lambda_s^s = \beta_s\Delta t_1 + \frac{\eta}{1-\theta}\Delta t_2, \\
\Delta\lambda_t^s &= \lambda_t^b - \lambda_t^s = -\frac{\eta}{\theta}\Delta t_2,
\end{aligned} \qquad (16)$$

$$\begin{aligned}
\Delta\mu_1^s &= \mu_1^b - \mu_1^s = \Delta t_1\left[(1-\theta) + \sqrt{\frac{\beta_s r_s\left(\frac{r_s}{\theta} + \frac{r_t}{1-\theta}\right)}{c_1}}\right] + \eta\Delta t_2, \\
\Delta\mu_2^s &= \mu_2^b - \mu_2^s = (1-\theta)\beta_s\Delta t_1.
\end{aligned} \qquad (17)$$

For more straightforward expression, we write optimal revenue $\pi^s = \pi^b - \Delta\pi^s$:

$$\Delta\pi^s = \Delta t_1^2 r_s\beta_s\sqrt{\frac{r_s\beta_s + \eta\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right)}{c_1}} + \Delta t_2\eta\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right), \qquad (18)$$

in which,

$$\Delta t_1 = \sqrt{\frac{c_1}{r_s\beta_s + \eta\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right)}} - \sqrt{\frac{c_1}{r_s\beta_s}},$$
$$\Delta t_2 = \sqrt{\frac{c_1}{r_s\beta_s + \eta\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right)}} - t. \tag{19}$$

It is really hard to confirm whether $\Delta\pi^s < 0$, since $\Delta\pi^s$ is affected by so many exogenous variables. Next we give some explanations for $\Delta\pi^s$ and give sufficient condition for avoiding revenue loss.

Δt_1 measures the difference of waiting time at stage 1 under inequity aversion and general condition, and $\Delta t_1 > 0$ if and only if $\theta r_s < (1-\theta)r_t$. Δt_2 measures the difference of waiting time between two channels. Both $\Delta\mu_1$ and $\Delta\mu_2$ are proportional to Δt_1 and Δt_2. Unlike the conclusion in Section 4.1, the total arrival rate changes $(1-\theta)\beta_s\Delta t_1$. It is easy to understand that, under inequity aversion to waiting time, there will be more sit-down consumers if consumers wait less for ordering. In the expression of $\Delta\pi^s$, it is easy to find that the first item is positive. Therefore, to make sure $\Delta\pi^s < 0$,there must be $\Delta t_2(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}) < 0$, restaurant earn less from consumers who wait longer. That is to say, by providing a certain degree of price subsidy, restaurant can avoid revenue loss caused by inequity aversion.

4.4. Asymmetry inequity aversion to time difference

As defined in Section 4.2, when consumers are more sensitive to waiting time and are free to leave the channel directly, inequity aversion results in loss of consumers inevitably. Then the optimization problem turns to:

$$\max_{\mu_1^a,\mu_2^a} \{p_s(1-\theta)\lambda_s^s + p_t\theta(1-\delta)\lambda_t^s - c_1\mu_1^a - c_2\mu_2^a\}$$
$$\text{s.t. } \lambda_s^s = \left[\lambda - \beta_s\left(\frac{1}{\mu_1^a - (1-\theta)\lambda_s^s} + \frac{1}{\mu_2^a - (1-\theta)\lambda_s^s - \theta\lambda_t^s}\right) - p_s + \frac{\eta}{1-\theta}\left(\min\left(\frac{1}{\mu_1^a - (1-\theta)\lambda_s^s}, t\right) - \frac{1}{\mu_1^a - (1-\theta)\lambda_s^s}\right)\right]^+,$$
$$\lambda_t^s = \left[\lambda - \beta_t\left(\frac{1}{\mu_2^a - (1-\theta)\lambda_s^s - \theta\lambda_t^s} + t\right) - p_t - c_d + \frac{\eta}{\theta}\left(\min\left(\frac{1}{\mu_1^a - (1-\theta)\lambda_s^s}, t\right) - t\right)\right]^+. \tag{20}$$

The optimal capacity level μ_1^a, μ_2^a and revenue π^a are given in Proposition 6.

Proposition 6. *When there is asymmetry inequity aversion to waiting time difference, the optimal capacity level and revenue need to be discussed in two scenarios:*

(1) *If* $\frac{1}{\mu_1^a - (1-\theta)\lambda_s^s} < t$. *we have*

$$\Delta\lambda_s^a = \lambda_s^b - \lambda_s^a = \beta_s\left(\sqrt{\frac{c_1}{r_s\beta_s - r_t\frac{\eta}{\theta}}} - \sqrt{\frac{c_1}{r_s\beta_s}}\right) > 0,$$
$$\Delta\lambda_t^a = \lambda_t^b - \lambda_t^a = \frac{\eta}{\theta}\left(\sqrt{\frac{c_1}{r_s\beta_s - r_t\frac{\eta}{\theta}}} - t\right) > 0. \tag{21}$$

$$\Delta\mu_1^a = \mu_1^b - \mu_1^a = (1-\theta)\Delta\lambda_s^a + \left(\sqrt{\frac{r_s\beta_s}{c_1}} - \sqrt{\frac{r_s\beta_s - r_t\frac{\eta}{\theta}}{c_1}}\right) > 0,$$
$$\Delta\mu_2^a = \mu_2^b - \mu_2^a = (1-\theta)\Delta\lambda_s^a + \theta\Delta\lambda_t^a > 0. \tag{22}$$

The total revenue changes

$$\Delta\pi^a = r_t\Delta\lambda_t^a + \left(r_s\beta_s\frac{c_1}{r_s\beta_s - r_t\frac{\eta}{\theta}} + \sqrt{c_1\left(r_s\beta_s - r_t\frac{\eta}{\theta}\right)} - 2\sqrt{c_1 r_s\beta_s}\right) > 0. \tag{23}$$

(2) *If* $\frac{1}{\mu_1^a-(1-\theta)\lambda_s^s} > t$. *we have*

$$\Delta\lambda_s^a = \lambda_s^b - \lambda_s^a = \beta_s\left(\sqrt{\frac{c_1}{r_s\left(\beta_s+\frac{\eta}{1-\theta}\right)}} - \sqrt{\frac{c_1}{r_s\beta_s}}\right) + \frac{\eta}{1-\theta}\left(\sqrt{\frac{c_1}{r_s\left(\beta_s+\frac{\eta}{1-\theta}\right)}} - t\right), \quad (24)$$
$$\Delta\lambda_t^a = \lambda_t^b - \lambda_t^a = 0.$$

$$\Delta\mu_1^a = (1-\theta)\Delta\lambda_s^a + \left(\sqrt{\frac{r_s\beta_s}{c_1}} - \sqrt{\frac{r_s\left(\beta_s+\frac{\eta}{1-\theta}\right)}{c_s}}\right), \quad (25)$$
$$\Delta\mu_2^a = (1-\theta)\Delta\lambda_s^a.$$

The total revenue changes

$$\begin{aligned}\Delta\pi^a &= 2\sqrt{c_1r_s}\left(\beta_s+\frac{\eta}{1-\theta}-\sqrt{\beta_s}\right) - \frac{\eta r_s t}{1-\theta}\\ &> \sqrt{\frac{c_1r_s}{\beta_s+\frac{\eta}{1-\theta}}}\left(\beta_s+\frac{\eta}{1-\theta}+\beta_s-2\sqrt{\beta_s\left(\beta_s+\frac{\eta}{1-\theta}\right)}\right)\\ &\geq 0.\end{aligned} \quad (26)$$

The results rely on time difference of two channels. Capacity levels and revenue changes differently in the two conditions.

When consumers spend more time on take-away channel, both arrival rates of two channels reduce. The arrival rate of sit-down channel reduce for longer waiting time, while the arrival rate of take-away channel reduce for inequity aversion. Then capacity levels in both stages decrease.

However, when consumers spend more time on sit-down channel, the results become more complicated. Arrival rate of take-away channel remains unchanged. And arrival rate of sit-down channel is influenced by inequity aversion and less waiting time at stage 1, where the former decreases it while the later increases it, making it hard to confirm whether $\Delta\lambda_s^a > 0$ or not. Therefore, the changes of capacity levels at two stages are also unknown. The only certain thing is the decreasing revenue.

4.5. Sensitivity analysis

In this section, we conduct a sensitivity analysis to understand how the key model parameters affect capacity level and revenue. For a more brief and easy-understanding result, we directly analyze the optimal solution obtained from the basic model. We consider how the two types of waiting time sensitivity β_i and commission ratio δ have various effects on the arrival rate λ_s, λ_t, the optimal capacity μ_1, μ_2 and the total operating income π. The analysis results are shown in Propositions 7–9.

Proposition 7. *The waiting time sensitivity of sit-down service have influence on the optimal solution as below:*

(1) *The arrival rate* λ_s *is decrease in* β_s*, while the arrival rate* λ_t *is independent of* β_s*.*
(2) *The total income* π *is decrease in* β_s*.*
(3) *Capacity level* μ_2 *is decrease in* β_s*. When* $p_s > 2c_1 + c_2$*, the capacity level* μ_1 *is a concave function of* β_s*, and when* $p_s \leq 2c_1 + c_2$*,* μ_1 *is decrease in* β_s*.*

First of all, we consider the impact of waiting time sensitivity of consumers who choose sit-down service. When β_s rises, the customers are less willing to wait and their arrival rate will drop, but since β_s does not change the kitchen lag, the arrival rate of consumers who choose take-away service is not affected, the total operating income is reduced. There are two possible conditions when we study the relationship between μ_1 and

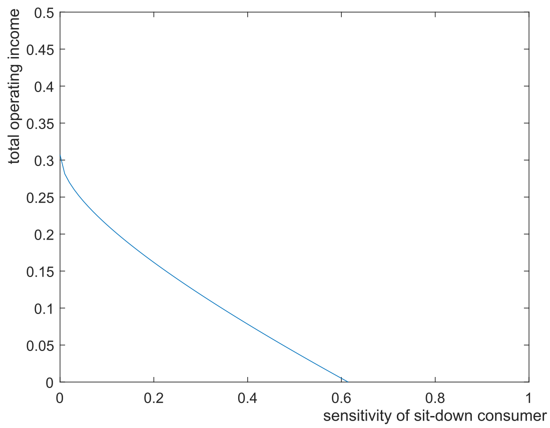

FIGURE 1. Impact of waiting time sensitivity β_s on total operating income π.

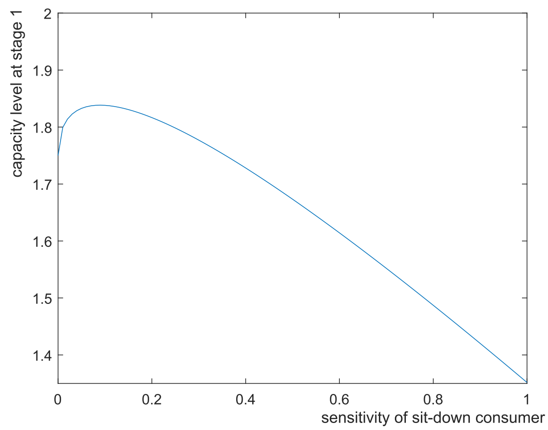

FIGURE 2. Impact of waiting time sensitivity β_s on front-end capacity level μ_1 (a) $c_1 = 0.1$.

β_s: if $\sqrt{\frac{r_s}{c_1}} > (1-\theta)\sqrt{\frac{c_1}{r_s}}$, *i.e.* $p_s > 2c_1 + c_2$, the front-end service cost is relatively low, so the safety capacity of the front-end service plays a more significant role. When β_s is small, μ_1 increases with β_s because the change of safety capacity dominates the change of μ_1 at this time. Since the front-end service cost is low, operator can fully respond to fluctuations in arrivals of shoppers by expanding the scale of front-end services. However, when β_s is large, the change of arrival rate λ_s, that is, the change of base capacity, dominates. At this time, due to the decrease in the number of customers, the operator should reduce the scale of front-end services. If $\sqrt{\frac{r_s}{c_1}} \leq (1-\theta)\sqrt{\frac{c_1}{r_s}}$, *i.e.* $p_s \leq 2c_1 + c_2$, the front-end service cost is relatively high and the impact of the base capacity is always dominant. When the demand for front-end capacity level increases but the operator cannot enhance the front-end capacity level to increase the total operating profit, the arrival rate of the consumers who prefer to sit-down service reduces, then the capacity level required at the front-end counter also decreases.

More intuitively, we conduct numerical stimulation with $\lambda = 3, \beta_t = 0.5, \theta = 0.3, t = 0.2, p_s = p_t = 0.5, c_d = 0.05, c_1 = 0.1, c_2 = 0.2, \delta = 0.2$. When β_s varies in $[0, 1]$, the variation of π is shown in Figure 1.

There are two cases of numerical simulation of the service level μ_1. Let $c_1 = 0.1$, at this time, $p_s = 0.5 > 2c_1 + c_2 = 0.4$, μ_1 varies as in Figure 2. And when $c_1 = 0.2$, $p_s = 0.5 \leq 2c_1 + c_2 = 0.6$, then we get Figure 3.

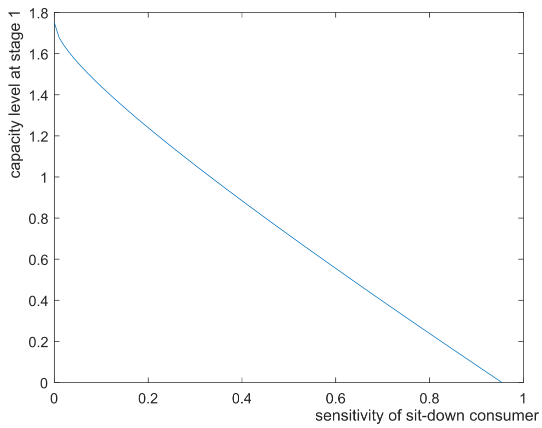

FIGURE 3. (b) $c_1 = 0.2$.

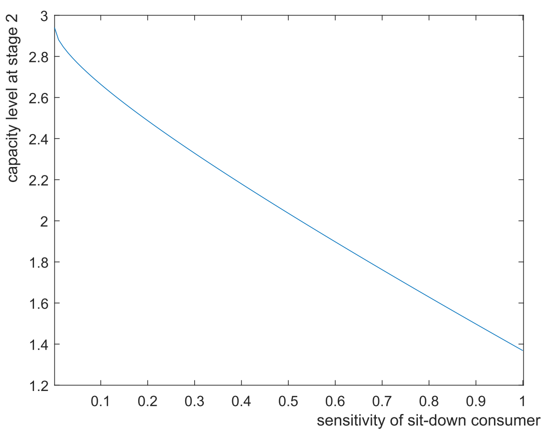

FIGURE 4. Impact of waiting time sensitivity β_s on front-end capacity level μ_2.

β_s affects μ_2 through λ_s mainly, so μ_2 is decremented with respect to β_s. The result of numerical stimulation is shown in Figure 4.

Similarly, we have the results about the waiting time sensitivity of consumers whose select take-away service:

Proposition 8. *The waiting time sensitivity of take-away service have influence on the optimal solution as below:*

(1) *The arrival rate λ_s is increase in β_t, while the arrival rate λ_t is decrease in β_t.*
(2) *The total operating income π is a concave function of β_t and reach the maximum at $\beta_t^* = \frac{\beta_s r_s}{r_s + 2r_t}$.*
(3) *Capacity level μ_1 is decrease in β_t. Capacity level μ_2 is a unimodal function of β_t, which first increases and then decreases.*

Consider the role of the waiting time sensitivity β_t. When β_t rises, the arrival rate λ_s will rise due to the decrease in kitchen lag, but consumers who select take-away service become less willing to wait so their arrival rate λ_t will drop. The total operating income π increases first and then decreases with the increase of β_t, and

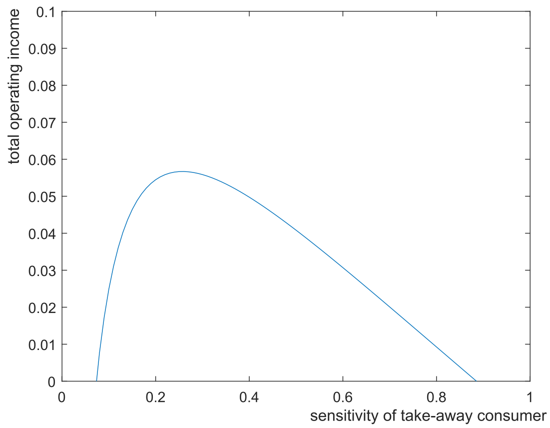

FIGURE 5. Impact of waiting time sensitivity β_t on total operating income π.

π take the maximum value at β_t^*. They have the relationship as follow:

$$\frac{\partial \pi}{\partial \beta_t} = r_s \frac{\partial \lambda_s}{\partial \beta_t} + r_t \frac{\partial \lambda_t}{\partial \beta_t} - \frac{1}{2}\sqrt{\frac{c_2(r_s+r_t)}{\beta_t}} = \frac{1}{2}\sqrt{\frac{c_2}{\beta_t(r_s+r_t)}}\left[\frac{r_s\beta_s}{\beta_t} - (r_s+2r_t)\right]. \tag{27}$$

When β_t is small, the arrival rate λ_s affects π dominantly, so π increases with the increase of β_t. When β_t is large, the arrival rate λ_t dominates, taking the place of λ_s, thus π decreases as β_t increases. We find that the optimal solution is $\beta_t^* = \frac{\beta_s r_s}{r_s+2r_t} = \frac{\beta_s}{1+2\frac{r_t}{r_s}}$. If r_t is relatively high, which means that take-away service brings more benefit, β_t^* is relatively small and β_t is more likely to reach the threshold that the arrival rate λ_t dominans, where take-away consumers are the major source of profits. β_t affects μ_1 through λ_s mainly, so μ_1 is incremented with respect to β_t. Whether $-\theta + \frac{r_s+r_t}{c_2}$ is greater than 0 or less than 0, μ_2 increases first and then decreases with regard to β_t. When β_t is small, the arrival rate λ_s predominates, so the kitchen's capacity level μ_2 increases with the increase of β_t. When λ_t is larger, λ_t surpasses λ_s to obtain the dominant influence position and the capacity level μ_2 decreases with the increase of λ_t.

Let $\beta_s = 0.5, \beta_t$ varies in [0,1], the result of numerical stimulation is shown in Figure 5.

β_t affects μ_1 through λ_s mainly, so μ_1 is incremented with respect to β_t. The result of numerical stimulation is shown in Figure 6.

Whether $-\theta + \frac{r_s+r_t}{c_2}$ is greater than 0 or less than 0, μ_2 increases first and then decreases with regard to β_t. When β_t is small, the arrival rate λ_s predominates, so the kitchen's capacity level μ_2 increases with the increase of β_t. When λ_t is larger, λ_t surpasses λ_s to obtain the dominant influence position and the capacity level μ_2 decreases with λ_t. The result of numerical stimulation is shown in Figure 7.

In general, waiting time sensitivity is meant to be negatively correlated with arrival rate in one channel, but may be positively correlated with the arrival rate of the other channel. When the sensitivity β is large enough, the total return π will eventually drop. The change of capacity level μ is more complicated, and different model parameters need to be analyzed separately.

Next, we discuss the effect of the commission ratio δ on the optimal solution. The analysis result is shown in the following proposition.

Proposition 9. *The commission ratio δ has the same effect on the customer arrival rate λ_s, λ_t, $\frac{\partial \lambda_s}{\partial \delta}, \frac{\partial \lambda_t}{\partial \delta} < 0$ and $\frac{\partial \lambda_s}{\partial \delta} : \frac{\partial \lambda_t}{\partial \delta} = \beta_s : \beta_t$; $\frac{\partial \mu_1}{\partial \delta}, \frac{\partial \mu_2}{\partial \delta} < 0$.*

Although the commission drawn by the third-party platform is only realized through the channel of take-out service, it can still affect the customers who choose sit-down service through influencing the capacity level of the

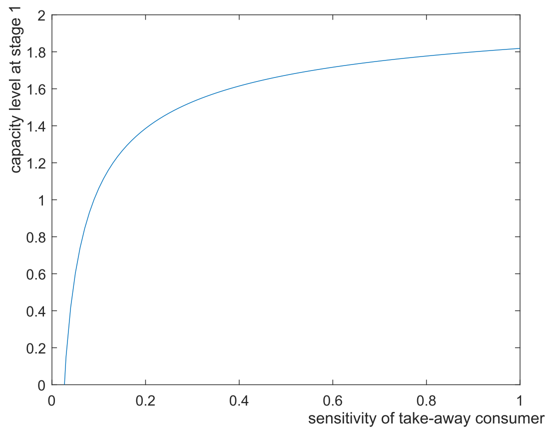

FIGURE 6. Impact of waiting time sensitivity β_t on counter capacity level μ_1.

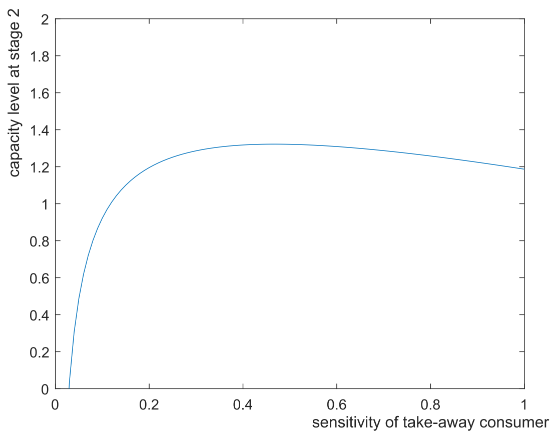

FIGURE 7. Impact of waiting time sensitivity β_t on kitchen capacity level μ_2.

kitchen, and the impact is only related to their waiting time sensitivity regardless of what channel they are in. In the expressions of λ_s, λ_t, δ just appears in the part of r_t, which is positively correlated with the kitchen lag, while the kitchen lag affects the arrival rate by waiting time sensitivity β. Therefore, λ decreases as δ increases, and when δ changes, the variation of λ_s, λ_t is proportional.

5. Additional discussions

In this section, we extend our model in two ways. For details, we study how to price in Section 5.1, and discuss the condition of self-delivery in Section 5.2.

5.1. Optimal price

One important assumption in our basic model is that price is determined by the market and restaurant has no pricing power. Now we extend the basic model by considering the price is also the decision variable and we add discussion about the optimal price of take-away channel and sit-down channel.

$$\max_{p_s,p_t,0\le(1-\theta)\lambda_s^b\le\mu_1,0\le(1-\theta)\lambda_s^b+\theta\lambda_t^b\le\mu_2} \left\{p_s(1-\theta)\lambda_s^b + p_t\theta(1-\delta)\lambda_t^b - c_1\mu_1 - c_2\mu_2\right\}$$
$$\text{s.t. } \lambda_s^b = \left[\lambda - \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s^b} + \frac{1}{\mu_2-(1-\theta)\lambda_s^b-\theta\lambda_t^b}\right) - p_s\right]^+, \quad (28)$$
$$\lambda_t^b = \left[\lambda - \beta_t\left(\frac{1}{\mu_2-(1-\theta)\lambda_s^b-\theta\lambda_t^b} + t\right) - p_t - c_d\right]^+.$$

Since we have known the optimal revenue under certain price, we study how revenue changes with respect to price in this section. As defined in above, p_s, p_t and r_s, r_t share a one-to-one mapping, we optimize $\pi(r_s, r_t)$ next, which is to find the optimal solution of a binary equation. By calculation we have,

$$\pi_s(r_s,r_t) = \frac{\partial\pi^b}{\partial r_s} = \lambda_s - \frac{r_s}{1-\theta}, \pi_t(r_s,r_t) = \frac{\partial\pi^b}{\partial r_t} = \lambda_s - \frac{r_t}{\theta}\cdot \quad (29)$$

Unfortunately, though we can solve a system of two element equations and get specific expressions of (r_s, r_t) in theory, it is hard to derive analytical solutions. We conduct some numerical analysis to π_t and π_t. π_s is a approximation of hook function of r_s, and $\pi_s < 0$ when $r_s \searrow 0$ or $r_s \nearrow \infty$. Therefore, if $\sup_{r_s,r_t}\pi_s \le 0$, π^b is a monotonic decreasing function of r_s, otherwise π_b decreases, increases and then decreases on r_s. So as the result of π^b and r_t. It should be noted that in Proposition 1, we give $\overline{\lambda}$ to make sure that for $\lambda > \overline{\lambda}$, restaurant keeps two channels, *i.e.* $r_s > 0, r_t > 0$. Under this circumstance, there must be (r_s, r_t) such that $\pi_s(r_s,r_t) > 0, \pi_t(r_s,r_t) > 0$.

Besides, taking $p_s = p_t$ to $\pi_s = 0, \pi_t = 0$, we find it does not always hold. So restaurant do have an incentive to set different price for two channels.

5.2. Self-Delivery

Most restaurants manage their take-away business by cooperating with the third party platform since their scale of operation cannot afford a complete delivery system. However, some large-scale chain restaurants, like MacDonald and KFC, do have their own delivery organization. At this time, they need to decide delivery time and bear the delivery cost. In this part, we add a decision variable t to the basic model. According to the reality, delivery cost $\psi(t)$ is a decreasing function of t, the cost approaches infinity when delivery time goes to 0, and decreases to 0 when delivery time is not restricted. Therefore, it is reasonable to assume that $\psi(t) = \frac{h}{t}$, where h is a positive constant. Then the optimization problem turns to:

$$\max_{\mu_1,\mu_2,t} \left\{p_s(1-\theta)\lambda_s^b + p_t\theta\lambda_t^b - c_1\mu_1 - c_2\mu_2 - \psi(t)\right\}$$
$$\text{s.t. } \lambda_s^b = \left[\lambda - \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s^b} + \frac{1}{\mu_2-(1-\theta)\lambda_s^b-\theta\lambda_t^b}\right) - p_s\right]^+, \quad (30)$$
$$\lambda_t^b = \left[\lambda - \beta_t\left(\frac{1}{\mu_2-(1-\theta)\lambda_s^b-\theta\lambda_t^b} + t\right) - p_t\right]^+.$$

It is easy to find that $t^* = \sqrt[3]{\frac{h}{(p_t-c_2)\beta_s}}$. Restaurant can choose the optimal delivery time based on unit net profit and waiting time sensitivity.

6. Conclusion

Capacity management and cost control have always been important contents in the decision-making of restaurants. Based on the assumption of long-term free market environment, we use a two-stage tandem queuing system to model the decision process, and discusses the change of the optimal solution by introducing inequity aversion.

We have the following conclusions:

(1) The optimal capacity level consists of two parts: a "base capacity" to satisfy the mean demand, and a "safety capacity" to hedges against variability in practical arrivals.
(2) The loss of operating income caused by waiting lag is equal to the waste arising from fluctuations in arrival rate. The total operating revenue can also be viewed as the mean revenue, which is brought by base capacity, subtracting the waste resulting from safety capacity, or operating loss caused by waiting lag.
(3) After the introduction of symmetry inequity aversion, the change in total operating income is determined by the relative unit service income, price difference and inefficiency factor between the two channels.
(4) After the introduction of asymmetry inequity aversion, the total operating income will reduce due to the loss of footfall.
(5) Waiting time sensitivity is meant to be negatively correlated with arrival rate in one channel, but may be positively correlated with arrival rate of the other channel. When the sensitivity β is large enough, the total return will eventually drop. With different model parameters, the change in capacity level varies. In the long run, the overall capacity level declines as the waiting time sensitivity of customers who choose sit-down service increases, and capacity level of kitchen declines as the waiting time sensitivity of customers who choose take-away service increases.
(6) The commissions ratio negatively affects the result through the channel of take-away service.

Our work has two main implications for restaurants: on one hand, it is necessary to conduct sufficient market research to realize the current market situation and correctly judge the relationship between each parameter (for example, the impact of waiting time sensitivity on capacity level depends on the relationships between cost and revenue, distribution of footfall and other factors); on the other hand, restaurants should pay attention to the waiting time sensitivity to different channels, and insensitive customers and sensitive customers could be priced separately.

In this paper, we aim at a free market that reaches equilibrium in the long term and construct a $M/M/1$ queueing system model. Based on the basic model, we further discuss the optimal pricing problem and self-delivery problem in the additional discussion. Furthermore, we can have some extended research for further research as follows. Firstly, compared with the situation in the long term that the market price is fixed, the operator adjusts the capacity level to change the operating cost, in the short term, the operator cannot immediately change the number of staff or work equipment and can only adjust the price, that is, he can change the price p instead of the service level μ. Secondly, in most cases, limited by restaurants' storefront area and the number of servers, the ability to serve customers is limited. The new customers observe the current system information such as the number of diners, the service ability, etc. before deciding whether to enter the restaurant, then it constitutes a $M/M/1/K$ queuing system observable by new entrants. However, the take-away service is not the case. The customer does not have to be limited by the capacity ability or the current waiting queue system, thus forming an $M/M/1$ queuing system. Although we build an elaborate optimization model, the catering market is much more complicated than the model. There is a Stackelberg game between the third-party platform and the operators and competition between operators. For restaurant operators who are price takers, this paper hopes to provide them with some long-term effective recommendations on cost control and capacity management. Thirdly, we should investigate a more realistic problem by focus on optimizing the number of services in the kitchen and the numbers of table/chairs in the restaurant.

Appendix A.

Proof of Proposition 1. The restaurant has three choices: (i) the restaurant chooses to keep the two channels. Consider the following optimization problem:

$$\max_{0\le(1-\theta)\lambda_s\le\mu_1,0\le(1-\theta)\lambda_s+\theta\lambda_t\le\mu_2} p_s(1-\theta)\lambda_s + p_t(1-\delta)\theta\lambda_t - c_1\mu_1 - c_2\mu_2$$

$$\text{s.t. } \lambda_s = \lambda - \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s} + \frac{1}{\mu_2-(1-\theta)\lambda_s-\theta\lambda_t}\right) - p_s, \qquad \text{(A.1)}$$

$$\lambda_t = \lambda - \beta_t\left(\frac{1}{\mu_2-(1-\theta)\lambda_s-\theta\lambda_t} + t\right) - p_t - c_d.$$

Because restaurant never pay cost more than the highest income, we have $c_1\mu_1+c_2\mu_2 \le p\lambda$, *i.e.* $\mu_1 \le \frac{p\lambda}{c_1}, \mu_2 \le \frac{p\lambda}{c_2}$. Thus, the constraint set $\{0 \le (1-\theta)\lambda_s \le \mu_1, 0 \le (1-\theta)\lambda_s + \theta\lambda_t \le \mu_2\}$ is boundary. Obviously, the set is closed, hence it is a compact set. Since the constraint set is compact and the objective function is continuous, by Weierstass Extreme Value Theorem [5], we can find a maximum. Then it is easy to check that if there is a positive solution $(\pi^b, \lambda_s^b, \lambda_t^b, \mu_1^b, \mu_2^b)$ of optimization problem (8), it must be a feasible solution and optimal solution to optimization problem (A.1), either.

The Lagrangian of the problem (A.1) is defined as follows:

$$\begin{aligned} L(\lambda_s,\lambda_t,\mu_1,\mu_2,\rho_1,\rho_2) &= p_s(1-\theta)\lambda_s + p_t(1-\delta)\theta\lambda_t - c_1\mu_1 - c_2\mu_2 \\ &\quad + \rho_1\left\{\lambda_s - \lambda + \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s} + \frac{1}{\mu_2-(1-\theta)\lambda_s-\theta\lambda_t}\right) + p_s\right\} \\ &\quad + \rho_2\left\{\lambda_t - \lambda + \beta_t\left(\frac{1}{\mu_2-(1-\theta)\lambda_s-\theta\lambda_t} + t\right) + p_t + c_d\right\} \end{aligned} \tag{A.2}$$

where $\rho_1, \rho_2 \in \Re$ are the Lagrange multipliers. To find the optimal solution, we solve the following equation set:

$$\begin{aligned} &\frac{\partial L}{\partial \lambda_s} = p_s(1-\theta) + \rho_1\left\{1 + \beta_s\frac{1-\theta}{[1-(1-\theta)\lambda_s]^2} + \beta_s\frac{1-\theta}{[\mu_2-(1-\theta)\lambda_s-\theta\lambda_t]^2}\right\} \\ &\qquad + \rho_2\beta_t\frac{1-\theta}{[\mu_2-(1-\theta)\lambda_s-\theta\lambda_t]^2} = 0, \\ &\frac{\partial L}{\partial \lambda_t} = p_t(1-\delta)\theta + \rho_2\left\{1+\beta_t\frac{\theta}{[\mu_2-(1-\theta)\lambda_s-\theta\lambda_t]^2}\right\} + \rho_1\beta_s\frac{\theta}{[\mu_2-(1-\theta)\lambda_s-\theta\lambda_t]^2} = 0, \\ &\frac{\partial L}{\partial \mu_1} = -c_1 - \frac{\rho_1\beta_s}{[1-(1-\theta)\lambda_s]^2} = 0, \\ &\frac{\partial L}{\partial \mu_2} = -c_2 - \frac{\rho_1\beta_s}{[\mu_2-(1-\theta)\lambda_s-\theta\lambda_t]^2} - \frac{\rho_2\beta_t}{[\mu_2-(1-\theta)\lambda_s-\theta\lambda_t]^2} = 0, \\ &\lambda_s - \lambda + \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s} + \frac{1}{\mu_2-(1-\theta)\lambda_s-\theta\lambda_t}\right) + p_s = 0, \\ &\lambda_t - \lambda + \beta_t\left(\frac{1}{\mu_2-(1-\theta)\lambda_s-\theta\lambda_t} + t\right) + p_t + c_d = 0, \end{aligned} \tag{A.3}$$

where μ_1, μ_2 satisfy: $\mu_1 \ge (1-\theta)\lambda_s, \mu_2 \ge (1-\theta)\lambda_s + \theta\lambda_t$.

Without the boundary conditions $\mu_1 - (1-\theta)\lambda_s \ge 0, \mu_2 - (1-\theta)\lambda_s - \theta\lambda_t \ge 0$, we can get that:

$$\begin{aligned} &\mu_1^b = (1-\theta)\lambda_s^b + \sqrt{\frac{\beta_s(1-\theta)(p_s-c_1-c_2)}{c_1}}, \\ &\mu_2^b = (1-\theta)\lambda_s^b + \theta\lambda_t^b + \sqrt{\frac{\beta_t[p_t(1-\delta)\theta + p_s(1-\theta) - c_1(1-\theta) - c_2]}{c_2}}, \\ &\lambda_s^b = \lambda - \beta_s\left(\sqrt{\frac{c_1}{\beta_s(1-\theta)(p_s-c_1-c_2)}} + \sqrt{\frac{c_2}{\beta_t[p_t(1-\delta)\theta + p_s(1-\theta) - c_1(1-\theta) - c_2]}}\right) - p_s, \\ &\lambda_t^b = \lambda - \beta_t\left(\sqrt{\frac{c_2}{\beta_t[p_t(1-\delta)\theta + p_s(1-\theta) - c_1(1-\theta) - c_2]}} + t\right) - p_t - c_d, \\ &\pi^b = r_s\lambda_s + r_t\lambda_t - \sqrt{c_1\beta_s r_s} - \sqrt{c_2\beta_t(r_s+r_t)}. \end{aligned} \tag{A.4}$$

So, if $0 \le (1-\theta)\lambda_s^b \le \mu_1^b, 0 \le (1-\theta)\lambda_s^b + \theta\lambda_t^b \le \mu_2^b$, $(\lambda_s^b, \lambda_t^b, \mu_1^b, \mu_2^b)$ must be an interior solution,then an optimal solution; otherwise, the optimal solution is a corner solution, then it is easy to check that the optimal value is nonnegative at this time and the restaurant never chooses this option. In order that $\lambda_s^b, \lambda_t^b, \pi^b > 0$, there must be:

$$\begin{aligned}
&\lambda > \beta_s\left(\sqrt{\frac{c_1}{\beta_s(1-\theta)(p_s-c_1-c_2)}} + \sqrt{\frac{c_2}{\beta_t[p_t(1-\delta)\theta+p_s(1-\theta)-c_1(1-\theta)-c_2]}}\right) + p_s,\\
&\lambda > \beta_t\left(\sqrt{\frac{c_2}{\beta_t[p_t(1-\delta)\theta+p_s(1-\theta)-c_1(1-\theta)-c_2]}} + t\right) + p_t + c_d, \qquad (A.5)\\
&\lambda > \frac{1}{r_1+r_2}\left[2\sqrt{c_1\beta_s r_s} + \sqrt{c_2\beta_s r_s} + \sqrt{c_2\beta_t(r_s+r_t)} + r_s p_s\right] + r_s\sqrt{\frac{c_2\beta_t}{r_s+r_t}} + r_t(p_t+c_d) + r_t\beta_t t.
\end{aligned}$$

Cited this condition as $\lambda > \overline{\lambda_1}$, with $\lambda > \overline{\lambda_1}$, $(\lambda_s^b, \lambda_t^b, \mu_s^b, \mu_t^b)$ become an interior and optimal solution to (A.1). Considering that $\lambda_s^b, \lambda_t^b > 0$, solution $(\lambda_s^b, \lambda_t^b, \mu_s^b, \mu_t^b)$ is also a feasible solution to optimization problem (8). Because problem (8) has more constraints than problem (A.1), its maximum is no larger than π^b. In conclusion, $(\lambda_s^b, \lambda_t^b, \mu_s^b, \mu_t^b)$ is an optimal solution to problem (8).

(ii) The restaurant chooses to keep the sit-down channel and close take-away channel. Then the problem can be reduced to:

$$\begin{aligned}
\max_{0\le(1-\theta)\lambda_s\le\mu_1,\mu_2}\ & p_s(1-\theta)\lambda_s - c_1\mu_1 - c_2\mu_2\\
\text{s.t. } & \lambda_s = \lambda - \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s} + \frac{1}{\mu_2-(1-\theta)\lambda_s}\right) - p_s.
\end{aligned} \qquad (A.6)$$

The Lagrangian is defined as follows:

$$\begin{aligned}
L(\lambda_s,\mu_1,\mu_2,\rho) = {} & p_s(1-\theta)\lambda_s - c_1\mu_1 - c_2\mu_2\\
& + \rho\left\{\lambda_s - \lambda + \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s} + \frac{1}{\mu_2-(1-\theta)\lambda_s}\right) + p_s\right\}
\end{aligned} \qquad (A.7)$$

where $\rho \in \Re$ is the Lagrange multiplier. To find the optimal solution, we solve the following equation set:

$$\begin{aligned}
p_s(1-\theta) + \rho + \rho\beta_s\frac{1-\theta}{(\mu_1-(1-\theta)\lambda_s)^2)} + \rho\beta_s\frac{1-\theta}{(\mu_2-(1-\theta)\lambda_s)^2)} &= 0\\
-c_1 - \rho\beta_s\frac{1}{(\mu_1-(1-\theta)\lambda_s)^2)} &= 0\\
-c_1 - \rho\beta_s\frac{1}{(\mu_1-(1-\theta)\lambda_s)^2)} &= 0\\
\lambda - \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s} + \frac{1}{\mu_2-(1-\theta)\lambda_s}\right) - p_s &= 0
\end{aligned} \qquad (A.8)$$

where μ_1, μ_2 satisfy $0 \le (1-\theta)\lambda_s \le \mu_1 < \frac{r_s\lambda}{c_1}, 0 \le (1-\theta)\lambda_s \le \mu_2 < \frac{r_s\lambda}{c_2}$.

The optimal solution is one of the critical points. Let's first ignore the constraints on μ_1, μ_2 and solve the equation set above, which gives us a unique solution:

$$\begin{aligned}
\lambda_s^* &= \lambda - \sqrt{\frac{c_1\beta_s}{r_s}} - \sqrt{\frac{c_2\beta_s}{r_s}} - p_s,\\
\mu_1^* &= (1-\theta)\lambda_s + \sqrt{\frac{\beta_s r_s}{c_1}},\\
\mu_2^* &= (1-\theta)\lambda_s + \sqrt{\frac{\beta_s r_s}{c_2}},\\
\pi^* &= r_s\lambda_s^* - \sqrt{c_1\beta_s r_s^*} - \sqrt{c_2\beta_s r_s^*}.
\end{aligned} \qquad (A.9)$$

Thus, if $0 \leq (1-\theta)\lambda_s \leq \mu_1 < \frac{r_s\lambda}{c_1}, 0 \leq (1-\theta)\lambda_s \leq \mu_2 < \frac{r_s\lambda}{c_2}$, $(\lambda_s^*, \mu_1^*, \mu_2^*)$ must be the optimal solution. If not, the optimal solution must be corner solution, *i.e.* $\lambda_s^* = 0$, $\mu_1^* = \frac{r_s\lambda}{c_1}$ or $\mu_2^* = \frac{r_s\lambda}{c_2}$, under which condition the optimal value is nonpositive, so the manager won't choose this sales program.

Therefore, $\pi^b > \pi^*$ if and only if:

$$\lambda > \overline{\lambda}_2 = \sqrt{\frac{\beta_t c_2}{r_1 + r_2}} + \beta_t t + p_t + c_d + \frac{\sqrt{c_2\beta_t(r_s + r_t)}}{r_t} + \frac{r_s\beta_s}{r_t}\sqrt{\frac{c_2}{\beta_t(r_s + r_t)}}.$$

(iii) The restaurant chooses to keep the take-away channel and close sit-down channel. Then the problem can be reduced to:

$$\begin{aligned} \max_{0\leq\theta\lambda_t\leq\mu_1,\mu_2} \quad & p_t(1-\delta)\theta\lambda_t - c_2\mu_2 \\ \text{s.t. } & \lambda_t = \lambda - \beta_t\left(\frac{1}{\mu_2 - \theta\lambda_t} + t\right) - p_t - c_d. \end{aligned} \tag{A.10}$$

The Lagrangian is defined as follows:

$$L(\lambda_t, \mu_2, \rho) = p_t(1-\delta)\theta\lambda_t - c_2\mu_2 + \rho\left\{\lambda_t - \lambda + \beta_t\left(\frac{1}{\mu_2 - (1-\theta)\lambda_s} + t\right) + p_t + c_d\right\} \tag{A.11}$$

where $\rho \in \Re$ is the Lagrange multiplier. To find the optimal solution, we solve the following equation set:

$$\begin{aligned} p_t(1-\delta)\theta + \rho + \rho\beta_t\frac{\theta}{(\mu_2 - \theta\lambda_t)^2} &= 0 \\ -c_2 - \rho\beta_s\frac{1}{(\mu_2 - \theta\lambda_t)^2)} &= 0 \\ \lambda_t - \lambda + \beta_t\frac{1}{\mu_2 - \theta\lambda_t} - p_t - c_d &= 0 \end{aligned} \tag{A.12}$$

where μ_2 satisfies $0 \leq \theta\lambda_t \leq \mu_2$.

The optimal solution is one of the critical points. Let's first ignore the constraints on μ_2 and solve the equation set above, which gives us a unique solution:

$$\begin{aligned} \lambda_t^* &= \lambda - \sqrt{\frac{c_2\beta_t}{r_t}} - p_t - c_d, \\ \mu_2^* &= \theta\lambda_t + \sqrt{\frac{\beta_t r_t}{c_2}}, \\ \pi^* &= r_t\lambda_t^* - \sqrt{c_2\beta_t r_t}. \end{aligned} \tag{A.13}$$

Thus, if $0 \leq \theta\lambda_t \leq \mu_2 < \frac{r_t\lambda}{c_2}$, (λ_t^*, μ_2^*) must be the optimal solution. If not, the optimal solution must be corner solution, *i.e.* $\lambda_s^* = 0$, or $\mu_2^* = \frac{r_s\lambda}{c_2}$, under which condition the optimal value is nonpositive, so the manager won't choose this sales program.

Therefore, $\pi^b > \pi_*$ if and only if:

$$0\lambda > \overline{\lambda}_3 = \sqrt{\frac{\beta_s c_1}{r_s}} + \beta_s\sqrt{\frac{c_2}{\beta_t(r_s + r_t)}} + p_s + \frac{r_t}{r_s}\sqrt{\frac{c_2\beta_s}{r_s + r_t}} + \sqrt{\frac{c_1\beta_s}{r_s}} + \frac{\sqrt{c_2\beta_t(r_s + r_t)}}{r_s} - 2\frac{\sqrt{c_2\beta_t r_t}}{r_s}.$$

We let $\overline{\lambda} = \max\{\overline{\lambda}_1, \overline{\lambda}_2, \overline{\lambda}_3\}$, when $\lambda > \overline{\lambda}$, we get the conclusion. □

Proof of Proposition 2. Because the waiting lag caused by offering service includes counter lag $\sqrt{\frac{c_1}{\beta_s(1-\theta)(p_s-c_1-c_2)}} = \sqrt{\frac{c_1}{\beta_s r_s}}$, which result from ordering at front-end counter, and kitchen lag, which arise from the preparation process, $\sqrt{\frac{c_2}{\beta_t[p_t(1-\delta)\theta+p_s(1-\theta)-c_1(1-\theta)-c_2]}} = \sqrt{\frac{c_2}{\beta_t(r_s+r_t)}}$. The product of waiting lag and latency sensitivity represents a decrease in consuming propensity resulting from waiting time, and the decrease in consuming propensity multiplied by unit service income is the loss of operating income due to waiting lag $\sqrt{c_1\beta_s r_s} + \sqrt{c_2\beta_t(r_s+r_t)}$, which is exactly the loss caused by fluctuation in arrival rate. The former is the loss arising from waiting lag, the latter is the loss arising from the spare time of counter, and they are exactly equal. □

Proof of Proposition 3. The basic optimization problem is:

$$\begin{aligned}
&\max_{\mu_1,\mu_2} \{p_s(1-\theta)\lambda_s^s + p_t\theta(1-\delta)\lambda_t^s - c_1\mu_1 - c_2\mu_2\}\\
&\text{s.t. } \lambda_s^s = \left[\lambda - \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s^s} + \frac{1}{\mu_2-(1-\theta)\lambda_s^s-\theta\lambda_t^s}\right) - p_s + \frac{\gamma}{1-\theta}(p_t-p_s)\right]^+,\\
&\qquad \lambda_t^s = \left[\lambda - \beta_t\left(\frac{1}{\mu_2-(1-\theta)\lambda_s^s-\theta\lambda_t^s} + t\right) - p_t - c_d + \frac{\gamma}{\theta}(p_s-p_t)\right]^+.
\end{aligned} \tag{A.14}$$

Compared to optimal problem (8), there are just some constant changing in problem (A.14), which exerts little difference to the solving process. Therefore, we know that optimal problem share similar optimal solution with (8). Same as the procedures in proof of Proposition 1, we have an optimal solution $\lambda_s^s, \lambda_t^s > 0$ when λ is large enough. In the following we just discuss the condition that λ is large enough to make sure arrival rate $\lambda_s^s, \lambda_t^s > 0$. For convenience, we denote $\Delta\lambda_s = \frac{\gamma}{1-\theta}(p_t-p_s), \Delta\lambda_t = \frac{\gamma}{\theta}(p_s-p_t)$. Given that the price p_s, p_t is fixed, the change of total income and the optimal capacity level is only in proportion to the change of arrival rate. Then from Proposition 1 we known:

$$\Delta\pi^s = r_s\Delta\lambda_s + r_t\Delta\lambda_t, \tag{A.15}$$

$$\Delta\mu_1^s = (1-\theta)\Delta\lambda_s, \quad \Delta\mu_2^s = \theta\Delta\lambda_t + (1-\theta)\Delta\lambda_s. \tag{A.16}$$

Take the expression of $\Delta\lambda_s, \Delta\lambda_t$ to (9) and (10), then we get the Proposition 3.

When there is symmetry inequity aversion, the total operating revenue π^s satisfies the equation

$$\pi^s = \pi^b + \Delta\pi^s = \pi^b + \frac{r_s\gamma}{1-\theta}(p_t-p_s) + \frac{r_t\gamma}{\theta}(p_s-p_t) = \pi^b + \gamma(p_t-p_s)\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right). \tag{A.17}$$

As above, the variation of operating income is determined by the difference of unit service income, price difference and inequity inefficiency factor between the two channels.

The capacity levels at two stages become

$$\mu_1^s = \mu_1^b + \Delta\mu_1^s = \mu_1^b + \gamma(p_t-p_s), \quad \mu_2^s = \mu_2^b + \Delta\mu_2^s = \mu_2^b. \tag{A.18}$$

The variation of capacity level depends simply on the amount of transferring. □

Proof of Proposition 4. For this asymmetry inequity aversion case, variation happens to the arrival rate λ_i:

$$(1-\theta)\lambda_s \Rightarrow (1-\theta)\lambda_s - \gamma[p_s - \min(p_s,p_t)] \tag{A.19}$$

$$\theta\lambda_t \Rightarrow \theta\lambda_t - \gamma[p_t - \min(p_s,p_t)]. \tag{A.20}$$

Compared to optimal problem (8), there are just some constant changing, which exerts little difference to the solving process. Therefore, we know that optimal problem share similar optimal solution with (8). Same as the procedures in proof of Proposition 1, we have an optimal solution $\lambda_s^a, \lambda_t^a > 0$ when λ is large enough. In the following we just discuss the condition that λ is large enough to make sure arrival rate $\lambda_s^a, \lambda_t^a > 0$. Let $\Delta\lambda_s^a = \frac{\gamma}{1-\theta}(\min(p_s,p_t) - p_s), \Delta\lambda_t^a = \frac{\gamma}{\theta}(\min(p_s,p_t) - p_t)$. For convenience, we just discuss the situation that the arrival rate $\lambda_s, \lambda_t > 0$. Then the optimization problem turns out to be:

$$\begin{aligned} &\max_{\mu_1,\mu_2} \{p_s(1-\theta)\lambda_s^a + p_t\theta(1-\delta)\lambda_t^a - c_1\mu_1 - c_2\mu_2\} \\ &\text{s.t. } \lambda_s^a = \left[\lambda - \beta_s\left(\frac{1}{\mu_1 - (1-\theta)\lambda_s^a} + \frac{1}{\mu_2 - (1-\theta)\lambda_s^a - \theta\lambda_t^a}\right) - p_s + \frac{\gamma}{1-\theta}(\min(p_s,p_t) - p_s)\right]^+, \\ &\lambda_t^a = \left[\lambda - \beta_t\left(\frac{1}{\mu_2 - (1-\theta)\lambda_s^a - \theta\lambda_t^a} + t\right) - p_t - c_d + \frac{\gamma}{\theta}(\min(p_s,p_t) - p_t)\right]^+. \end{aligned} \quad \text{(A.21)}$$

The same as the analysis in Section 4.1, we suppose that the price p_s, p_t is fixed, the change of total income π and capacity levels is only relative to the change of arrival rate λ, and we get that there is asymmetry inequity aversion, the change of operating revenue $\Delta\pi^a$ follows the equation

$$\pi^a = \pi^b + \Delta\pi^a = \pi^b + r_s\frac{\gamma}{1-\theta}(\min(p_s,p_t) - p_s) + r_t\frac{\gamma}{\theta}(\min(p_s,p_t) - p_t) < \pi^b. \quad \text{(A.22)}$$

As for optimal capacity levels μ_1, μ_2, we have similar results:

$$\mu_1^a = \mu_1^b + \Delta\mu_1^a = \mu_1^b + \gamma(\min(p_s,p_t) - p_s), \quad \text{(A.23)}$$

$$\mu_2^a = \mu_2^b + \Delta\mu_2^a = \mu_2^b + \gamma(\min(p_s,p_t) - p_s) + \gamma(\min(p_s,p_t) - p_t) = \mu_2^b - \gamma|p_s - p_t|. \quad \text{(A.24)}$$

□

Proof of Proposition 5. The optimization problem turns to be:

$$\begin{aligned} &\max_{\mu_1,\mu_2} \{p_s(1-\theta)\lambda_s^s + p_t\theta(1-\delta)\lambda_t^s - c_1\mu_1 - c_2\mu_2\} \\ &\text{s.t. } \lambda_s^s = \left[\lambda - \beta_s\left(\frac{1}{\mu_1 - (1-\theta)\lambda_s^s} + \frac{1}{\mu_2 - (1-\theta)\lambda_s^s - \theta\lambda_t^s}\right) - p_s + \frac{\eta}{1-\theta}\left(t - \frac{1}{\mu_1 - (1-\theta)\lambda_s^s}\right)\right]^+, \\ &\lambda_t^s = \left[\lambda - \beta_t\left(\frac{1}{\mu_2 - (1-\theta)\lambda_s^s - \theta\lambda_t^s} + t\right) - p_t - c_d + \frac{\eta}{\theta}\left(\frac{1}{\mu_1 - (1-\theta)\lambda_s^s} - t\right)\right]^+. \end{aligned} \quad \text{(A.25)}$$

Same as the procedures in proof of Proposition 1, we have an optimal solution $\lambda_s^s, \lambda_t^s > 0$ when λ is large enough. In the following we just discuss the condition that λ is large enough to make sure arrival rate $\lambda_s^s, \lambda_t^s > 0$. The Lagrangian of the problem above defined:

$$\begin{aligned} &L(\lambda_s, \lambda_t, \mu_1, \mu_2, \rho_1, \rho_2) \\ &= p_s(1-\theta)\lambda_s + p_t(1-\delta)\theta\lambda_t - c_1\mu_1 - c_2\mu_2 \\ &+ \rho_1\left\{\lambda_s^s - \lambda + \beta_s\left(\frac{1}{\mu_1 - (1-\theta)\lambda_s^s} + \frac{1}{\mu_2 - (1-\theta)\lambda_s^s - \theta\lambda_t^s}\right) + p_s - \frac{\eta}{1-\theta}\left(t - \frac{1}{\mu_1 - (1-\theta)\lambda_s^s}\right)\right\} \\ &+ \rho_2\left\{\lambda_t^s - \lambda + \beta_t\left(\frac{1}{\mu_2 - (1-\theta)\lambda_s^s - \theta\lambda_t^s} + t\right) + p_t + c_d - \frac{\eta}{\theta}\left(\frac{1}{\mu_1 - (1-\theta)\lambda_s^s} - t\right)\right\} \end{aligned} \quad \text{(A.26)}$$

where $\rho_1, \rho_2 \in \Re$ are the Lagrange multipliers. To find the optimal solution, we solve the following equation set:

$$
\begin{aligned}
\frac{\partial L}{\partial \lambda_s^s} &= p_s(1-\theta) + \rho_1\left\{1 + \left(\beta_s + \frac{\eta}{1-\theta}\right)\frac{1-\theta}{[\mu_1-(1-\theta)\lambda_s^s]^2} + \beta_s\frac{1-\theta}{[\mu_2-(1-\theta)\lambda_s^s-\theta\lambda_t^s]^2}\right\}\\
&\quad + \rho_2\left\{-\frac{\eta}{\theta}\frac{1-\theta}{[\mu_1-(1-\theta)\lambda_s^s]^2} + \beta_t\frac{1-\theta}{[\mu_2-(1-\theta)\lambda_s^s-\theta\lambda_t^s]^2}\right\} = 0,\\
\frac{\partial L}{\partial \lambda_t^s} &= p_t\theta(1-\delta) + \rho_1\beta_s\frac{\theta}{[\mu_2-(1-\theta)\lambda_s^s-\theta\lambda_t^s]^2} + \rho_2\left\{1 + \beta_t\frac{\theta}{[\mu_2-(1-\theta)\lambda_s^s-\theta\lambda_t^s]^2}\right\} = 0,\\
\frac{\partial L}{\partial \mu_1} &= -c_1 - \frac{\rho_1\left(\beta_s + \frac{\eta}{1-\theta}\right)}{[\mu_1-(1-\theta)\lambda_s^s]^2} + \frac{\rho_2\frac{\eta}{\theta}}{[\mu_2-(1-\theta)\lambda_s^s-\theta\lambda_t^s]^2} = 0,\\
\frac{\partial L}{\partial \mu_2} &= -c_2 - \frac{\rho_1\beta_s + \rho_2\beta_t}{[\mu_2-(1-\theta)\lambda_s^s-\theta\lambda_t^s]^2} = 0,\\
\lambda_s^s - \lambda &+ \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s^s} + \frac{1}{\mu_2-(1-\theta)\lambda_s^s-\theta\lambda_t^s}\right) + p_s - \frac{\eta}{1-\theta}\left(t - \frac{1}{\mu_1-(1-\theta)\lambda_s^s}\right) = 0,\\
\lambda_t^s - \lambda &+ \beta_t\left(\frac{1}{\mu_2-(1-\theta)\lambda_s^s-\theta\lambda_t^s} + t\right) + p_t + c_d - \frac{\eta}{\theta}\left(\frac{1}{\mu_1-(1-\theta)\lambda_s^s} - t\right) = 0.
\end{aligned}
\tag{A.27}
$$

Same as the proof of Proposition 1, We have $\rho_1 = (1-\theta)(c_1+c_2-p_s) = -r_s, \rho_2 = \theta[c_2-(1-\delta)p_t] = -r_t$ and

$$
\lambda_s^s = \lambda - \left(\beta_s + \frac{\eta}{1-\theta}\right)\sqrt{\frac{c_1}{\eta\left(\frac{r_s}{1-\theta}\right) + r_s\beta_s}} - \beta_s\sqrt{\frac{c_2}{r_s\beta_s + r_t\beta_t}} + \frac{\eta t}{1-\theta} - p_s, \tag{A.28}
$$

$$
\lambda_t^s = \lambda + \frac{\eta}{\beta}\sqrt{\frac{c_1}{\eta(\frac{r_s}{1-\theta}) + r_s\beta_s}} - \beta_s\sqrt{\frac{c_2}{r_s\beta_s + r_t\beta_t}} - \left(\beta_t + \frac{\eta}{\theta}\right)t - p_t - c_d. \tag{A.29}
$$

Given λ_s^b, λ_t^b, we get

$$
\begin{aligned}
\Delta\lambda_s^s &= \lambda_s^b - \lambda_s^s = \beta_s\Delta t_1 + \frac{\eta}{1-\theta}\Delta t_2,\\
\Delta\lambda_t^s &= \lambda_t^b - \lambda_t^s = -\frac{\eta}{\theta}\Delta t_2,
\end{aligned}
\tag{A.30}
$$

where $\Delta t_1 = \sqrt{\frac{c_1}{r_s\beta_s + \eta\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right)}} - \sqrt{\frac{c_1}{r_s\beta_s}}, \Delta t_2 = \sqrt{\frac{c_1}{r_s\beta_s + \eta\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right)}} - t$. Taking them to expression of μ_1^s, μ_2^s, and π^s:

$$
\begin{aligned}
\mu_1^b - \mu_1^s &= (1-\theta)\left(\lambda_s^b - \lambda_s^s\right) - \sqrt{\frac{r_s\left(\beta_s + \frac{\eta}{1-\theta}\right) - r_t\frac{\eta}{\theta}}{c_1}} + \sqrt{\frac{r_s\beta_s}{c_1}}\\
&= \Delta t_1\left[(1-\theta) + \sqrt{\frac{\beta_s r_s\left(\frac{r_s}{\theta} + \frac{r_t}{1-\theta}\right)}{c_1}}\right] + \eta\Delta t_2,\\
\mu_2^b - \mu_2^s &= (1-\theta)\left(\lambda_s^b - \lambda_s^s\right) + \theta\left(\lambda_t^b - \lambda_t^s\right)\\
&= (1-\theta)\beta_s\Delta t_1,
\end{aligned}
\tag{A.31}
$$

$$
\begin{aligned}
\pi^b - \pi^s &= r_s\big(\lambda_s^b - \lambda_s^s\big) + r_t\big(\lambda_t^b - \lambda_t^s\big) + \sqrt{c_1\beta_s r_s + \eta c_1\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right)} - \sqrt{c_1\beta_s r_s} \\
&= r_s\left(\beta_s \Delta t_1 + \frac{\eta}{1-\theta}\Delta t_2\right) - r_t\frac{\eta}{\theta}\Delta t_2 - \Delta t_1\sqrt{\beta_s r_s\left[\beta_s r_s + \eta\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right)\right]} \\
&= \Delta t_1\sqrt{r_s\beta_s}\left[\sqrt{r_s\beta_s} - \sqrt{\beta_s r_s + \eta\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right)}\right] + \eta\Delta t_2\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right) \\
&= \Delta t_1^2 r_s\beta_s\sqrt{\frac{r_s\beta_s + \eta\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right)}{c_1}} + \Delta t_2\eta\left(\frac{r_s}{1-\theta} - \frac{r_t}{\theta}\right).
\end{aligned}
\tag{A.32}
$$

□

Proof of Proposition 6. Same as the procedures in proof of Proposition 1, we have an optimal solution $\lambda_s^a, \lambda_t^a > 0$ when λ is large enough. In the following we just discuss the condition that λ is large enough to make sure arrival rate $\lambda_s^a, \lambda_t^a > 0$. Next we discuss in two cases:

(i) $\frac{1}{\mu_1-(1-\theta)\lambda_s^s} < t$. Consumers in take-away channel wait longer than those who are in sit-down channel. Then we have optimal problem:

$$
\begin{aligned}
&\max_{\mu_1,\mu_2}\ \{p_s(1-\theta)\lambda_s^a + p_t\theta(1-\delta)\lambda_t^a - c_1\mu_1 - c_2\mu_2\} \\
&\text{s.t. } \lambda_s^a = \left[\lambda - \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s^a} + \frac{1}{\mu_2-(1-\theta)\lambda_s^a - \theta\lambda_t^a}\right) - p_s\right]^+, \\
&\qquad \lambda_t^a = \left[\lambda - \beta_t\left(\frac{1}{\mu_2-(1-\theta)\lambda_s^a - \theta\lambda_t^a} + t\right) - p_t - c_d + \frac{\eta}{\theta}\left(\frac{1}{\mu_1-(1-\theta)\lambda_s^a} - t\right)\right]^+.
\end{aligned}
\tag{A.33}
$$

Solving the equation set:

$$
\begin{aligned}
\frac{\partial L}{\partial\lambda_s^a} &= p_s(1-\theta) + \rho_1\left\{1 + \beta_s\frac{1-\theta}{[\mu_1-(1-\theta)\lambda_s^a]^2} + \beta_s\frac{1-\theta}{[\mu_2-(1-\theta)\lambda_s^a-\theta\lambda_t^a]^2}\right\} \\
&\quad + \rho_2\left\{-\frac{\eta}{\theta}\frac{1-\theta}{[\mu_1-(1-\theta)\lambda_s^a]^2} + \beta_t\frac{1-\theta}{[\mu_2-(1-\theta)\lambda_s^a-\theta\lambda_t^a]^2}\right\} = 0, \\
\frac{\partial L}{\partial\lambda_t^s} &= p_t\theta(1-\delta) + \rho_1\beta_s\frac{\theta}{[\mu_2-(1-\theta)\lambda_s^a-\theta\lambda_t^a]^2} + \rho_2\left\{1 + \beta_t\frac{\theta}{[\mu_2-(1-\theta)\lambda_s^a-\theta\lambda_t^a]^2}\right\} = 0, \\
\frac{\partial L}{\partial\mu_1^s} &= -c_1 - \frac{\rho_1\beta_s}{[\mu_1-(1-\theta)\lambda_s^a]^2} + \frac{\rho_2\frac{\eta}{\theta}}{[\mu_2-(1-\theta)\lambda_s^a-\theta\lambda_t^a]^2} = 0, \\
\frac{\partial L}{\partial\mu_2^s} &= -c_2 - \frac{\rho_1\beta_s + \rho_2\beta_t}{[\mu_2-(1-\theta)\lambda_s^a-\theta\lambda_t^a]^2} = 0, \\
&\lambda_s^a - \lambda + \beta_s\left(\frac{1}{\mu_1-(1-\theta)\lambda_s^a} + \frac{1}{\mu_2-(1-\theta)\lambda_s^a-\theta\lambda_t^a}\right) + p_s = 0, \\
&\lambda_t^a - \lambda + \beta_t\left(\frac{1}{\mu_2-(1-\theta)\lambda_s^a-\theta\lambda_t^a} + t\right) + p_t + c_d - \frac{\eta}{\theta}\left(\frac{1}{\mu_1-(1-\theta)\lambda_s^a} - t\right) = 0.
\end{aligned}
\tag{A.34}
$$

We have:

$$\begin{aligned} \lambda_s^a &= \lambda - \beta_s\left(\sqrt{\frac{c_1}{r_s\beta_s - r_t\frac{\eta}{\theta}}} + \sqrt{\frac{c_2}{r_s\beta_s + r_t\beta_t}}\right) - p_s < \lambda_s^b, \\ \lambda_t^a &= \lambda - \beta_t\left(\sqrt{\frac{c_2}{r_s\beta_s + r_t\beta_t}} + t\right) - p_t - c_d + \frac{\eta}{\theta}\left(\sqrt{\frac{c_1}{r_s\beta_s - r_t\frac{\eta}{\theta}}} - t\right) < \lambda_t^b. \end{aligned} \tag{A.35}$$

Taking them in the expressions of μ_1^a, μ_2^a and π^a, and we have:

$$\begin{aligned} \mu_1^b - \mu_1^a &= (1-\theta)\big(\lambda_s^b - \lambda_s^a\big) + \left(\sqrt{\frac{r_s\beta_s}{c_1}} - \sqrt{\frac{r_s\beta_s - r_t\frac{\eta}{\theta}}{c_1}}\right) > 0, \\ \mu_2^b - \mu_2^a &= (1-\theta)\Delta\lambda_s^a + \theta\Delta\lambda_t^a > 0. \end{aligned} \tag{A.36}$$

$$\begin{aligned} \pi^b - \pi^a &= r_s(\lambda_s^b - \lambda_s^a) + r_t(\lambda_t^b - \lambda_t^a) + \sqrt{c_1(r_s\beta_s - r_t\frac{\eta}{\theta})} - \sqrt{c_1\beta_s r_s} \\ &= r_t\Delta\lambda_t^a + \left[r_s\beta_s\sqrt{\frac{c_1}{r_s\beta_s - r_t\frac{\eta}{\theta}}} + \sqrt{c_1\left(r_s\beta_s - r_t\frac{\eta}{\theta}\right)} - 2\sqrt{c_1 r_s\beta_s}\right]. \end{aligned} \tag{A.37}$$

According to Jensen's inequality, the second item in $\pi^b - \pi^a$ is positive, so $\pi^b - \pi^a > 0$.

(ii) $\frac{1}{\mu_1 - (1-\theta)\lambda_s^a} > t$. Consumers in sit-down channel wait longer than those who are in take-away channel. Then we have optimal problem:

$$\begin{aligned} &\max_{\mu_1,\mu_2} \{p_s(1-\theta)\lambda_s^a + p_t\theta(1-\delta)\lambda_t^a - c_1\mu_1 - c_2\mu_2\} \\ &\text{s.t. } \lambda_s^a = \left[\lambda - \beta_s\left(\frac{1}{\mu_1 - (1-\theta)\lambda_s^a} + \frac{1}{\mu_2 - (1-\theta)\lambda_s^a - \theta\lambda_t^a}\right) - p_s + \frac{\eta}{1-\theta}\left(t - \frac{1}{\mu_1 - (1-\theta)\lambda_s^a}\right)\right]^+, \\ &\quad\;\; \lambda_t^a = \left[\lambda - \beta_t\left(\frac{1}{\mu_2 - (1-\theta)\lambda_s^a - \theta\lambda_t^a} + t\right) - p_t - c_d\right]^+. \end{aligned} \tag{A.38}$$

Solving the equation set:

$$\begin{aligned} \frac{\partial L}{\partial \lambda_s^a} &= p_s(1-\theta) + \rho_1\left\{1 + \left(\beta_s + \frac{\eta}{1-\theta}\right)\frac{1-\theta}{[\mu_1 - (1-\theta)\lambda_a^s]^2} + \beta_s\frac{1-\theta}{[\mu_2 - (1-\theta)\lambda_s^a - \theta\lambda_t^a]^2}\right\} \\ &\quad + \rho_2\beta_t\frac{1-\theta}{[\mu_2 - (1-\theta)\lambda_s^a - \theta\lambda_t^a]^2} = 0, \\ \frac{\partial L}{\partial \lambda_t^a} &= p_t\theta(1-\delta) + \rho_1\beta_s\frac{\theta}{[\mu_2 - (1-\theta)\lambda_s^a - \theta\lambda_t^a]^2} + \rho_2\left\{1 + \beta_t\frac{\theta}{[\mu_2 - (1-\theta)\lambda_s^a - \theta\lambda_t^a]^2}\right\} = 0, \\ \frac{\partial L}{\partial \mu_1^a} &= -c_1 - \frac{\rho_1\left(\beta_s + \frac{\eta}{1-\theta}\right)}{\mu_1 - (1-\theta)\lambda_s^s]^2} = 0, \\ \frac{\partial L}{\partial \mu_2^a} &= -c_2 - \frac{\rho_1\beta_s + \rho_2\beta_t}{[\mu_2 - (1-\theta)\lambda_s^a - \theta\lambda_t^a]^2} = 0, \\ \lambda_s^a &- \lambda + \beta_s\left(\frac{1}{\mu_1 - (1-\theta)\lambda_s^a} + \frac{1}{\mu_2 - (1-\theta)\lambda_s^a - \theta\lambda_t^a}\right) + p_s - \frac{\eta}{1-\theta}\left(t - \frac{1}{\mu_1 - (1-\theta)\lambda_s^a}\right) = 0, \\ \lambda_t^a &- \lambda + \beta_t\left(\frac{1}{\mu_2 - (1-\theta)\lambda_s^a - \theta\lambda_t^a} + t\right) + p_t + c_d = 0. \end{aligned} \tag{A.39}$$

We have:

$$\lambda_s^a = \lambda - \sqrt{\frac{c_1[(1-\theta)\beta_s+\eta]}{r_s}} - \beta_s\sqrt{\frac{c_2}{r_s\beta_s+r_t\beta_t}} - p_s + \frac{\eta t}{1-\theta},$$
$$\lambda_t^a = \lambda - \beta_t\left(\sqrt{\frac{c_2}{r_s\beta_s+r_t\beta_t}} + t\right) - p_t - c_d = \lambda_t^b. \tag{A.40}$$

Taking them in the expressions of μ_1^a, μ_2^a and π^a:

$$\mu_1^b - \mu_1^a = (1-\theta)\Delta\lambda_s^a + \left(\sqrt{\frac{r_s\beta_s}{c_1}} - \sqrt{\frac{r_s(\beta_s+\frac{\eta}{1-\theta})}{c_s}}\right), \tag{A.41}$$
$$\mu_2^b - \mu_2^a = (1-\theta)\Delta\lambda_s^a.$$

$$\begin{aligned}\pi^b - \pi^a &= r_s(\lambda_s^b - \lambda_s^a) + \sqrt{c_1 r_s\left(\beta_s + \frac{\eta}{1-\theta}\right)} - \sqrt{c_1 r_s\beta_s}\\ &= 2\sqrt{c_1 r_s}\left(\sqrt{\beta_s+\frac{\eta}{1-\theta}} - \sqrt{\beta_s}\right) - \frac{\eta r_s t}{1-\theta}.\end{aligned} \tag{A.42}$$

Since $\frac{1}{\mu_1-(1-\theta)\lambda_s^s} > t$, by Jensen's inequality, we have

$$\begin{aligned}\pi^b - \pi^a &> \sqrt{\frac{c_1 r_s}{\beta_s+\frac{\eta}{1-\theta}}}\left(\beta_s + \frac{\eta}{1-\theta} + \beta_s - 2\sqrt{\beta_s\left(\beta_s+\frac{\eta}{1-\theta}\right)}\right)\\ &\geq 0.\end{aligned} \tag{A.43}$$

□

Proof. From formulae (3) to (7), we know that:

$$\frac{\partial\lambda_s^b}{\partial\beta_s} = -\sqrt{\frac{c_2}{\beta_t(r_s+r_t)}} - \frac{1}{2}\sqrt{\frac{c_1}{\beta_s r_s}} < 0, \tag{A.44}$$

$$\frac{\partial\lambda_t^b}{\partial\beta_s} = 0, \tag{A.45}$$

$$\frac{\partial\pi^b}{\partial\beta_s} = r_s\frac{\partial\lambda_s^b}{\partial\beta_s} + r_t\frac{\partial\lambda_t^b}{\partial\beta_s} - \frac{1}{2}\sqrt{\frac{c_1 r_t}{\beta_s}} = -r_s\sqrt{\frac{c_2}{\beta_t(r_s+r_t)}} - \sqrt{\frac{c_1}{\beta_s r_s}} < 0, \tag{A.46}$$

$$\frac{\partial\mu_1^b}{\partial\beta_s} = (1-\theta)\frac{\partial\lambda_s^b}{\partial\beta_s} + \frac{1}{2}\sqrt{\frac{r_s}{\beta_s c_1}} = \frac{1}{2\sqrt{\beta_s}}[\sqrt{r_s}c_1 - (1-\theta)\sqrt{c_1}r_s] + (\theta-1)\sqrt{\frac{c_2}{\beta_t(r_s+r_t)}}, \tag{A.47}$$

where $\frac{1}{2}\sqrt{\frac{r_s}{\beta_s c_1}}$ represents the influences of safety capacity, $(1-\theta)\frac{\partial\lambda_s^b}{\partial\beta_s}$ represents the influences of base capacity. When $\frac{r_s}{c_1} < (1-\theta)\frac{c_1}{r_s}$, we always have $\frac{\partial\mu_1^b}{\beta_s} < 0$. When $\frac{r_s}{c_1} > (1-\theta)\frac{c_1}{r_s}$, let $\beta_s^* = \frac{\beta_t(r_s+r_t)[r_s-(1-\theta)c_1]^2}{4(1-\theta)^2 c_1 c_2 r_s}$, then $\frac{\partial\mu_1^b}{\beta_s} > 0$ if $\beta_s < \beta_s^*$, $\frac{\partial\mu_1^b}{\beta_s} < 0$ if $\beta_s > \beta_s^*$, so μ_1 is an upper convex function about β_s, which increases first and then decreases.

$$\frac{\partial\mu_2^b}{\partial\beta_s} = (1-\theta)\frac{\partial\lambda_s^b}{\partial\beta_s} + \theta\frac{\partial\lambda_t^b}{\partial\beta_s} = (1-\theta)\frac{\partial\lambda_s^b}{\partial\beta_s} < 0. \tag{A.48}$$

□

Proof. From formulae (3) to (7), we know that:

$$\frac{\partial \lambda_s^b}{\partial \beta_t} = \frac{\beta_s}{2\beta_t}\sqrt{\frac{c_2}{\beta_t(r_s+r_t)}} > 0, \tag{A.49}$$

$$\frac{\partial \lambda_t^b}{\partial \beta_t} = -\frac{1}{2}\sqrt{\frac{c_2}{\beta_t(r_s+r_t)}} < 0, \tag{A.50}$$

$$\frac{\partial \pi^b}{\partial \beta_t} = r_s\frac{\partial \lambda_s^b}{\partial \beta_t} + r_t\frac{\partial \lambda_t^b}{\partial \beta_t} - \frac{1}{2}\sqrt{\frac{c_2}{(r_s+r_t)}} = \frac{1}{2}\sqrt{\frac{c_2}{\beta_t(r_s+r_t)}}\left[\frac{r_t\beta_s}{\beta_t} - (r_s+2r_t)\right]. \tag{A.51}$$

When $\frac{\partial \pi^b}{\partial \beta_t} > 0$, $\frac{\partial \lambda_s}{\partial \beta_t}$ takes a major role in the above formula. Let $\beta_t^* = \frac{r_s\beta_s}{r_s+2r_t} = \frac{\beta_t}{1+2\frac{r_t}{r_s}}$, then $\frac{\partial \pi^b}{\partial \beta_t} > 0$ if $\beta_t < \beta_t^*$, $\frac{\partial \pi^b}{\partial \beta_t} < 0$ if $\beta_t > \beta_t^*$, that is, π increases first and then decreases with the increase of β_t and reaches the maximum at β_t^*.

$$\frac{\partial \mu_1^b}{\partial \beta_t} = (1-\theta)\frac{\partial \lambda_s^b}{\partial \beta_t} > 0, \tag{A.52}$$

$$\begin{aligned}\frac{\partial \mu_2^b}{\partial \beta_t} &= (1-\theta)\frac{\partial \lambda_s^b}{\partial \beta_t} + \theta\frac{\partial \lambda_t^b}{\partial \beta_t} + \frac{1}{2}\sqrt{\frac{r_s+r_t}{\beta_t c_2}} \\ &= \frac{1}{2}\sqrt{\frac{c_2}{\beta_t(r_s+r_t)}}\left[\frac{(1-\theta)\beta_s}{\beta_t} - \theta + \frac{r_s+r_t}{c_2}\right] - t\theta.\end{aligned} \tag{A.53}$$

When $\frac{\partial \mu_2^b}{\partial \beta_t} > 0$, $\frac{\partial \lambda_t^b}{\partial \beta_t}$ takes a major role in the above formula. μ_2 is also incremented first and then decremented with respect to β_t but it is not a convex function. □

Proof. From formulae (3) to (7), we know that:

$$\frac{\partial \lambda_s^b}{\partial \delta} = -\frac{1}{2}\beta_s p_t\theta\sqrt{\frac{c_2}{\beta_t}}(r_s+r_t)^{-\frac{3}{2}} < 0, \tag{A.54}$$

$$\frac{\partial \lambda_t^b}{\partial \delta} = -\frac{1}{2}\beta_t p_t\theta\sqrt{\frac{c_2}{\beta_t}}(r_s+r_t)^{-\frac{3}{2}} = \frac{\beta_t}{\beta_s}\frac{\partial \lambda_s^b}{\partial \delta} < 0, \tag{A.55}$$

$$\frac{\partial \mu_1^b}{\partial \delta} = (1-\theta)\frac{\partial \lambda_s^b}{\partial \delta} < 0, \tag{A.56}$$

$$\frac{\partial \mu_2^b}{\partial \delta} = (1-\theta)\frac{\partial \lambda_s^b}{\partial \delta} + \theta\frac{\partial \lambda_t^b}{\partial \delta} < 0. \tag{A.57}$$

As a result, the commission ratio draws down the operating income by reducing the arrival rate of two channels proportionally. □

Acknowledgements

This paper is supported by the National Natural Science Foundation of China (NSFC) under Grant Nos. 72261147707 and 72132007.

References

[1] J. Acimovic and S.C. Graves, Mitigating spillover in online retailing via replenishment. *Manuf. Serv. Oper. Manag.* **19** (2017) 419–436.

[2] A. Ackere, C. Haxholdt and E.R. Larsen, Dynamic capacity adjustments with reactive customers. *Omega* **41** (2013) 689–705.

[3] H. Ahn, I. Duenyas and M.E. Lewis, Optimal control of a two-stage tandem queueing system with flexible servers. *Probab. Eng. Inf. Sci.* **16** (2002) 453–469.

[4] A.S. Albana, Y. Frein and R. Hammami, Effect of a lead time-dependent cost on lead time quotation, pricing, and capacity decisions in a stochastic make-to-order system with endogenous demand. *Int. J. Prod. Econ.* **203** (2018) 83–95.

[5] T. Andreescu, C. Mortici and M. Tetiva, The extreme value theorem, in Mathematical Bridges. Birkhäuser, New York, NY (2017).

[6] A. Bassamboo, R.S. Randhawa and A. Zeevi, Capacity sizing under parameter uncertainty: safety staffing principles revisited. *Manag. Sci.* **56** (2010) 1668–1686.

[7] D.R. Bell, S. Gallino and A. Moreno, Inventory showrooms and customer migration in omni-channel retail: the effect of product information. *SSRN Electron. J.* (2013). DOI: 10.2139/ssrn.2370535.

[8] G.E. Bolton and A. Ockenfels, ERC: a theory of equity, reciprocity, and competition. *Am. Econ. Rev.* **90** (2000) 166–193.

[9] L.E. Bolton, L. Warlop and J.W. Alba, Consumer perceptions of price fairness. *J. Consum. Res.* **29** (2003) 474–491.

[10] D. Campbell and F. Frei, Cost structure, customer profitability, and retention implications of self-service distribution channels: evidence from customer behavior in an online banking channel. *Manag. Sci.* **56** (2010) 4–24.

[11] Y. Chen and T.H. Cui, The benefit of uniform price for branded variants. *Mark. Sci.* **32** (2013) 36–50.

[12] T.H. Cui, J.S. Raju and Z.J. Zhang, Fairness and channel coordiantion. *Manag. Sci.* **53** (2007) 1303–1314.

[13] C. Ding, K.H. Wang and R. Ran, Marketing channel coordination machanismbased on fairness perference. *J. Manag. Sci. China* **16** (2013) 80–94.

[14] E. Fehr and K.M. Schmidt, A theory of fairness, competition and cooperation. *Q. J. Econ.* **114** (1999) 817–868.

[15] F. Gao and X. Su, Online and offline information for omnichannel retailing. *Manuf. Serv. Oper. Manag.* **19** (2017) 84–98.

[16] F. Gao and X. Su, Omnichannel service operations with online and offline self-order technologies. *Manag. Sci.* **64** (2018) 3595–3608.

[17] L. Guo, Inequity aversion and fair selling. *J. Mark. Res.* **52** (2013) 77–89.

[18] J.M. Hall, P.K. Kopalle and D.F. Pyke, Static and dynamic pricing of excess capacity in a make-to-order environment. *Prod. Oper. Manag.* **18** (2009) 411–425.

[19] P. Harsha, S. Subramanian and J. Uichanco, Dynamic pricing of omnichannel inventories. *Manuf. Serv. Oper. Manag.* **21**(2019) 47–65.

[20] M. Haviv and R.S. Randhawa, Pricing in queues without demand information. *Manuf. Serv. Oper. Manag.* **16** (2014) 329–341.

[21] C.Y. Heo, S. Lee, A. Mattila and C. Hu, Restaurant revenue management: do perceived capacity scarcity and price differences matter? *Int. J. Hosp. Manag.* **35** (2013) 316–326.

[22] T.H. Ho and X.M. Su, Peer-induced fairness in games. *Am. Econ. Rev.* **99** (2009) 2022–2049.

[23] S.H. Hum, M. Parlar and Y. Zhou, Measurement and optimization of responsiveness in supply chain networks with queueing structures. *Eur. J. Oper. Res.* **264** (2018) 106–118.

[24] J. Hwang, L. Gao and W. Jang, Joint demand and capacity management in a restaurant system. *Eur. J. Oper. Res.* **207** (2010) 465–472.

[25] S. Kimes and J. Collier, How customers view self-service technologies. *MIT Sloan Manag.* **57** (2015) 25–26.

[26] J. Kurz, Capacity planning for a maintenance service provider with advanced information. *Eur. J. Oper. Res.* **251** (2016) 466–477.

[27] C. Lee and A.R. Ward, Optimal pricing and capacity sizing for the GI/GI/1 queue. *Oper. Res. Lett.* **42** (2014) 527–531.

[28] J.P. Medhi, Stochastic Models in Queueing Theory. Academic Press, New York, NY (1991).

[29] Z.Y. Niu, P. Huang and J. Wang, Online-offline uniform pricing strategy of multi-channel retailer based onfaieness preference. *Chin. J. Manag. Sci.* **25** (2017) 147–155.

[30] I. Papachristos and D.J. Pandelis, Optimal server assignment in a two-stage tandem queueing system. *Oper. Res. Lett.* **48** (2020) 71–77.

[31] E. Walster, E. Berscheid and W. Walster, New directions in equity research. *J. Pers. Soc. Psychol.* **25** (1973) 151–176.

[32] Y. Xiong, W. Yan, K. Fernandes, Z.K. Xiong and N. Guo, Clicks versus Bricks: the role of durability in marketing channel strategy of durable goods manufacturers. *Eur. J. Oper. Res.* **265** (2018) 909–918.

[33] Z.W. Zhan, M.H. Jiang and X.P. Wang, The optimal capacity decision of the catering merchant in omnichannel-service, production and delivery capacity. *Kybernetes* **53** (2023) 1958–1986.

[34] D.Z. Zhao and J.Yang, Research on decision-making modelof estimated delivery time in O2O food delivery. *Ind. Eng. Manag.* **23** (2018) 8–14.