OPTIMAL PRICING AND ORDERING STRATEGIES FOR DUAL-CHANNEL RETAILING WITH DIFFERENT SHIPPING POLICIES

NING LI AND ZHENG WANG∗
School of Automation, Southeast University
Nanjing, Jiangsu 210096, China

(Communicated by Ada Che)

ABSTRACT. In this paper, considering dual-channel retailing (online channel and offline channel), we study the pricing and ordering problem under different shipping policies. In this research, we mainly consider three shipping policies: without shipping price (OSP), with shipping price (WSP) and conditional free shipping (CFP). Based on the principle of maximum utility, we firstly obtain the probability of demand for the online and offline channels and further model the pricing and ordering problem under the three shipping policies. Further, avoiding the curse of dimensionality, the deep deterministic policy gradient (DDPG) method is employed to solve the problem to obtain the optimal pricing and ordering policy. Finally, we conduct some numerical experiments to compare the optimal pricing and ordering quantity under the three different shipping policies and reveal some managerial insights. The results show that the conditional free shipping policy is better than the other two policies, and stimulates the increase of demand to gain more profit.

1. Introduction. Recently, with the development of Internet technology, traditional retailers have begun to adopt dual-channel retailing (i.e., online channel and offline channel) in order to provide extensive choices for customers and gain more profit [23]. Some large offline retailers, such as Nike, IBM and Apple have constructed online retailing channel, while some online retailers such as Amazon, JD.com and Suning.com have also built brick-and-mortar store to increase the convenience of the local customers and expand the demand of products [28].

From the perspective of customers, the development of e-retail leads to the diversification of consumption habits, which further requires traditional retailers to improve the retailing channel. Some customers choosing online channel can reduce the time of traveling to physical store and searching products at physical store. However, they usually need to afford the delivery time. The customers choosing offline channel will enjoy better service in the physical store and get products immediately without waiting for delivery. However, they need to spend time on walking to the physical store and searching for the products. Therefore, by combining online

2020 Mathematics Subject Classification. Primary: 90B05, 90C15, 90C39.
Key words and phrases. Pricing, ordering, dual-channel, shipping price, deep deterministic policy gradient method.
This research is supported by National Natural Science Foundation of China under grant No. 61673109.
∗ Corresponding author: Zheng Wang.
channel and offline channel, the dual-channel retailing can satisfy multiple types of consumption habit and gradually become a mainstream model.

Notably, comparing with the online and offline channels, one of the salient characteristics of the difference of those channels is that the products from the online channel are directly delivered by the retailer, while the products from the offline channel are purchased when the customers arrive at the physical store. Therefore, the online channel will generate extra cost (i.e., shipping cost). A proper shipping policy is beneficial to the retailers and influences the customers’ purchasing behaviors [32]. In addition, the dual-channel retailers need to research the impact of different shipping policies on the dual-channel retailing.

Although customers and retailers can obtain more benefits from dual-channel retailing, the pricing and inventory management problem of a dual-channel retailing system becomes more difficult. For example, comparing with traditional offline retailing, dual-channel retailing needs to consider competitive demand between the online and the offline channels. Besides, in practice, the demands of product in the online and offline channels are stochastic depending on the price of products and the shipping price. Therefore, the pricing and inventory management policy for dual-channel retailing under different shipping policies becomes more difficult and more sophisticated.

Based on the above discussion, for satisfying the diversification of consumption habits and attracting more customers, we research the inventory management and pricing problem for dual-channel retailing, in which customers can purchase products from online channel and offline channel. Further, aiming at the shipping pricing of online channel, we research the problem under different shipping policies. Our research can provide the guideline for dual-channel retailers to employ different shipping policies to obtain more profit.

In this paper, we investigate the optimal pricing and ordering problems for dual-channel retailing under different shipping policies. The dual-channel retailer sells products through both online and offline channels. For the shipping cost, three types of shipping policies are considered: (1) without shipping price (OSP): the retailers afford the shipping price and the customers choosing online channel do not need afford the shipping price; (2) with shipping price (WSP): the customers choosing online channel need to afford the shipping price, while the retailers do not afford it; and (3) conditional free shipping (CFP): the customers who buy one product from online channel need to afford the shipping price, while they also can buy two products to avoid affording the shipping price. Based on the three shipping policies, we first construct the utility functions of a customer for online channel and offline channel under the three shipping policies. By the principle of maximum utility, we can obtain the probabilities of a customer’s choices for the online and offline channels, and further construct the probability of demand for the online and offline channels under the three different shipping policies to model the pricing and ordering problems. Finally, to avoid the curse of dimensionality, we employ the deep deterministic policy gradient method (DDPG) to solve those problems to obtain the optimal pricing and ordering strategies. Based on the proposed optimal ordering and pricing policy under different shipping policies, we can find that the shipping price policies have significant influence on the ordering and pricing policy for dual-channel retailing. Between the WSP policy and the OSP policy, the dual-channel retailer can employ the OSP policy for the lower shipping price and employ the WSP policy for the higher shipping price. In addition, the CFP policy can stimulate the
customers to purchase more product, and obtain more profit compared with other policies (i.e., the WSP policy and the OSP policy), no matter what the shipping price is high or low.

The rest of the paper is organized as follows. Section 2 reviews relevant literature. In Section 3, we describe the problems under different shipping polices and present the utility functions of online and offline channels under these polices, respectively. In Section 4, the probability of demand for the online and offline channels under different shipping policies is constructed and the pricing and ordering problems under these policies are formulated. In Section 5, we employ the constructed models to obtain the corresponding Bellman equations for the different shipping policies, respectively. Then the DDPG method is used to solve those Bellman equations to obtain the optimal pricing and ordering strategy. In Section 6, numerical experiments are conducted to evaluate the proposed optimal pricing and ordering strategy under different shipping policies and reveal some managerial insights. Section 7 concludes the paper and presents some issues for future research.

2. Literature review. With the rapid development of e-commerce, many retailers have begun to extend the sales channel. For example, traditional retailers began to set online sale channel, and large online retailers also constructed offline sale channel. Therefore, a lot of research have been paid attention to this trend. Chiang et al. (2003)[5] considered the online channel as a direct sale channel and found that it can increase the total profit significantly. Huang and Swaminathan (2009)[14] studied an optimal pricing problem for dual-channel retailing (online and offline channel) and proposed an optimal pricing policy for the monopoly and the duopoly case, respectively. Based on the Fleischer’s model, Ding et al. (2010)[6] researched the Bahncard problem considering fluctuated Bahncard price for online traveler and employ the customers’ tolerance for the risk to obtain the optimal risk algorithm. Tang et al. (2021)[27] employed the Stackelberg game method to propose a novel coordinating contract based on credit term and designed a sharing profit strategy to incent the cooperation between manufacturers and retailers under the dual-channel retailing.

For the dual-channel retailing, there is a rapid-growing stream of literature on dynamic pricing and inventory management [2, 7, 29]. Hua et al. (2010)[12] modeled two kinds of dual-channel supply chains (i.e., centralized and decentralized supply chain), and further studied the optimal pricing problem with delivery lead time, which indicated that the delivery lead time has a significant influence on the optimal pricing policy and the profit. Considering the quality of product, Chen et al. (2017)[4] developed an optimal pricing and quality investment policy for three scenarios (i.e., offline channel, online channel and dual channels) respectively. Li et al. (2019)[19] proposed an optimal pricing and service effort policy for dual-channel supply chain. Besides, the authors analyzed the impact of the showrooming effect on this policy. Considering the refund in dual-channel retailing, Li et al. (2019)[18] proposed a return strategy and a pricing policy, and found that the using of the return strategy on a dual-channel supply chain depends on the value of customer’s return rate. Li et al. (2015)[20] constructed a demand model depending on the inventory level and further proposed a heuristic method to research the inventory management problem for dual-channel retailing. He et al. (2018)[11] considered product’s perishability under the dual-channel supply chain and proposed the optimal pricing, inventory management and ordering frequency policy for centralized...
and decentralized models under the dual-channel retailing. Ryan et al. (2013)\cite{25} modelled a pricing and inventory management problem for a dual-channel supply chain and employed the Nash game and the Stackelberg game methods to obtain two different supply chain contracts for different situations. Considering the information asymmetry, Liu et al. (2010)\cite{22} proposed a joint optimal pricing and inventory management policy for online dual channel supply chain system under centralized control and decentralized control.

In practice, the demand of products is usually stochastic, which adds difficulties to the pricing and ordering problems for dual-channel retailing. In [1], an optimal inventory control policy for the online and offline channels was investigated, considering stochastic demand and random delivery lead time. Under known distribution function of stochastic demand, Yao et al. (2005)\cite{30} researched the impact of information sharing on inventory management and buy-back pricing policy for dual-channel supply chain. Huang et al. (2018)\cite{13} employed the Stackelberg game approach to investigate a joint ordering and pricing policy for a dual-channel supply chain including one retailer and one manufacturer, and analyzed the impact of the stochastic demand fluctuation on this policy. Roy et al. (2016)\cite{24} modeled a two-echelon dual-channel supply chain, and developed an optimal pricing, inventory management and service effort policy by using the Stackelberg game approach.

Geunes and Su (2020)\cite{9} took uncertain demand and space limitation into account to obtain the two-stage stochastic optimal model in order to determine the optimal product assortment, pricing and inventory level for dual-channel retailing. Li et al. (2014)\cite{17} modeled a Nash bargaining model for a dual-channel retailing with a risk-averse retailer and a risk-averse manufacturer, where the stochastic demand followed a uniform distribution. Based on this model, the authors determined the pricing and ordering policy.

Importantly, it is necessary to consider how to process the shipping cost for online channel, in order to increase more profits. Zhang et al (2019)\cite{31} mold an inventory management problem for the online retailing considering a shipping fee to the discrete newsvendor problem and employed Weak Aggregating Algorithm to solve this problem in order to obtain the optimal ordering policy. Gümüs et al. (2013)\cite{10} researched two pricing partitioning formats based on whether the price of product included the shipping and handling fees, and employed a two-stage game-theory approach to decide the pricing policy and pricing partitioning format. Becerril-Arreola et al. (2013)\cite{3} employed Arena to simulate the online retailing supply chain system and further used OptQuest to obtain the optimal pricing, free-shipping threshold and ordering policy. Huang and Cheng (2015)\cite{15} researched the impact of the threshold free shipping policies (piece-based and dollar-based) on the consumers’ shopping intention under the online retailing, and found that the policy based on piece stimulates more customer’s shopping desire than the policy based on dollar. Shao (2017)\cite{26} considered the calculated shipping policy and the free shipping policy, and further modeled the pricing competition and Cournot competition for online retailing. Based on the model, he employed the geographical pricing method and the uniform pricing method to obtain the optimal pricing and ordering policy.

From the above analysis, we can find that related literatures on the shipping policy [3, 10, 15, 26, 31] emphasize the online retailing that customer can only purchase product from online channel, without considering the competition between online channel and offline channel. Besides, the literatures [10, 15, 31] only research
the inventory management or pricing policy and the literatures [15, 26] only research the problem about shipping policy under the deterministic demand. For dual-channel retailing, although related literatures on dual-channel retailing [2, 4, 7, 11, 12, 19, 18, 20, 22, 25, 29] consider the competition between the online channel and the offline channel, they ignore the impact of shipping price on the online channel.

Therefore, we can find that there is no research on the optimal pricing and ordering problem for the dual-channel retailing considering different shipping policies and the competition between online channel and offline channel. Motivated by this research gap, our research contributes to existing literatures in the following aspects: (1) We construct the utility functions for the online channel and the offline channel under three different shipping policies and employ the principle of maximum utility to obtain the probability of demand for the two channels, which describes the competition between them under different shipping policies; (2) Considering three types of shipping policies (i.e., OSP, WSP and CFP), we research the pricing and ordering problem for the dual-channel retailing and compare the impact of the different shipping policies on the pricing and ordering strategies. Especially comparing with the related literature [31], we research the joint optimal inventory and pricing policy for dual-channel retailing considering the competition between the online and offline channels, while the literature [31] only researches the inventory management for online retailing without considering the competition. In addition, we research three shipping policies, while the literature [31] only considers the free-shipping policy. For avoiding the curse of dimensionality and reducing computation time, we employ the DDPG method to solve the pricing and ordering problem to obtain the joint optimal pricing and ordering policy under the different shipping policies.

3. Problem description. We consider a dual-channel retailer who sells a single type of products via online and offline channels. Therefore, the customers can purchase the products from the two channels. We assume that: (1) the brick-and-mortar store can provide online channel and offline channel; (2) each customer purchases at most one product by online or offline channel without considering shipping price [8]; (3) the dual-channel retailers employ a consistent pricing strategy and therefore the price $p_k$ denotes the selling price for both channels in the ordering period $k$ [8, 17], and the selling price is greater than shipping price $p_s$ (i.e., $p_k > p_s$); (4) the total amount of customers is $m$ in each ordering period; (5) shortage of the products is allowed and backlogged; (6) the customers from the offline channel bring the cross-selling profit $r$ [16]. It is noted that the customers from offline channel can evaluate the products in reality and are easily influenced by the recommendation from the clerks at the store, while the online customers usually have a clear purpose for products and pay more attention on shipping price. Compared with the cross-selling profit in online channel, the offline cross-selling profit is much greater. Therefore, we only consider the offline cross-selling profit, which is consistent with [16].

For the online channel, the retailers usually need to consider how to set up the shipping price in order to maximize the profit. In this paper, we consider three shipping policies (i.e., OSP, WSP and CFP). Under the OSP policy, the retailers afford the shipping price, while the customers from online channel afford the shipping price under the WSP policy. In practice, a lot of dual-channel retailers employ the conditional free shipping policy to stimulate customers to purchase
more products in order to obtain more profit. Generally, the customers who buy one product from the online channel need to afford the shipping price, while they can buy two products to avoid affording the shipping price under the CFP policy. Figure 1 shows the dual-channel retailing for the three shipping policies.

Figure 1. Dual-channel retailing for different shipping policies: (a) OSP; (b) WSP; (c) CFP

Based on the consumer utility theory and the above analysis, we can obtain the utility function of a customer for offline channel and the utility function of a customer for online channel under the three shipping policies. The customer choosing offline channel needs to go to the store and search the products (i.e., the offline purchasing cost $h_s$). Therefore, the utility function of the customer for offline channel (i.e., $u_s$) is

$$u_s = v - p^k - h_s$$

where $v$ is the value of a product.

Under the OSP policy (i.e., without shipping price), the customer from the online channel needs to search and pay for the products online, and wait for products to arrive, which generates the online purchasing cost $h_o$. Therefore, the utility function of the customer for online channel under the OSP policy is

$$u_{OSP} = v - p^k - h_o$$

Under the WSP policy (i.e., with shipping price), the customer from the online channel also needs to consider the searching and paying cost for online channel and the waiting for products to arrive (i.e., the online purchasing cost $h_o$). Besides, he must afford the shipping price $p_s$. Therefore, the utility function of the customer from the online channel under the WSP policy is

$$u_{WSP} = v - p^k - p_s - h_o$$

Under the CFP policy (i.e., conditional free shipping), the online customer also needs to consider the searching and paying cost for online channel and the waiting for products to arrive (i.e., the online purchasing cost $h_o$). Importantly, he can purchase two products to avoid affording the shipping price. However, it is noted that the utility of the customer for the excess product is discounted. Therefore, the
total online utility function for purchasing two products under the CFP policy can be described by the following equation:

\[ u_{CFP}^2 = (1 + \theta)v - 2p^k - h_o \]  

where \( \theta \) is the discount parameter of customer’s utility for the second product.

For comparing with the other utility functions under different shipping policies (i.e., \( u_s \), \( u_{osp} \) and \( u_{wsp} \)) to obtain the probabilities of a customer’s choices for the online and offline channels, the utility function of the customer from the online channel under the CFP policy is the average of the total online utility function for purchasing two products under the CFP policy, which can be described by the following equation:

\[ u_{CFP} = \frac{u_{CFP}^2}{2} = \frac{1 + \theta}{2}v - p^k - \frac{h_o}{2} \]  

Based on the principle of maximum utility, the customer can decide how to choose the online channel or the offline channel and whether to purchase more products to avoid affording the shipping price, under different shipping policies.

4. The pricing and ordering problems under different shipping policies.

In this section, we employ the constructed utility functions of a customer for dual-channel purchasing under different shipping policies to obtain the probabilities of the demands from online and offline channels. Based on these probabilities of those demands, we present the pricing and ordering problems under different shipping policies.

4.1. The pricing and ordering problem under the OSP policy.

In this situation, the customers can choose online channel and offline channel to purchase products at the selling price \( p^k \). The customers from the online channel do not need to consider the shipping price \( p_s \), since the dual-channel retailer affords the shipping price. Therefore, we obtain the utility functions of an individual customer under the OSP policy: \( u_s = v - p^k - h_s \) (i.e., offline utility function under the OSP policy) and \( u_{OSP} = v - p^k - h_o \) (i.e., online utility function under the OSP policy), where the purchasing costs of the customer from the two channels (i.e., \( h_o \) and \( h_s \)) are in the set \( [0, H] \times [0, H] \), where \( H > v - p^k \geq 0 \).

Based on the principle of maximum utility, the customer purchases product through the channel with greater utility. When these utilities (i.e, \( u_s \) and \( u_{OSP} \)) are less than zero, the customer will not purchase any products. Therefore, in this situation, a customer has three choices depending on different online purchasing cost \( h_o \) and offline purchasing cost \( h_s \), as shown in Figure 2. From Figure 2, the probability that a customer chooses the online channel under the OSP policy is the ratio of the area of “online” to the total area (i.e., \( H \times H \)). Similarly, the probability that a customer chooses the offline channel is the ratio of the area of “offline” to the total area (i.e., \( H \times H \)). Further, we can obtain those probabilities under the OSP policy as below:

\[ D^{OSP}_s(p^k) = \frac{1}{2H^2} (2H - v + p^k)(v - p^k) \]  

\[ D^{OSP}_o(p^k) = \frac{1}{2H^2} (2H - v + p^k)(v - p^k) \]
Figure 2. The customer’s choices for purchasing channels under the OSP policy

where $D_{OSP}^s$ is the probability that the customer chooses the offline channel and $D_{OSP}^o$ is the probability that the customer chooses the online channel under the OSP policy. From Eqs. (6) and (7), we can obtain the following proposition:

**Proposition 1.** Under the OSP policy, the probabilities of the customer’s choice satisfy:

1. The probability of the customer’s choice for the online channel (i.e., $D_{OSP}^o$) is equal to the probability of the customer’s choice for the offline channel (i.e., $D_{OSP}^s$) under the OSP policy.
2. With the increase of the price of product, the probabilities of the customer’s choice for the online channel and the offline channel (i.e., $D_{OSP}^o$ and $D_{OSP}^s$) decrease.

**Proof.** The correctness of this proposition can be directly observed from Eqs. (6) and (7).

From Proposition 1, we can find that the price of product has significant influence on the probabilities of the customer’s choice, while those probabilities are not affected by the shipping price under the OSP policy. Further, the probability of demand for the online and offline channels under the OSP policy can be obtained as below:

\[
\Psi_{OSP}(x_o^k, x_s^k, p^k) = \frac{m!}{x_o^k!x_s^k!(m - x_o^k - x_s^k)!} (D_{OSP}^o(p^k))^{x_o^k} \\
(D_{OSP}^s(p^k))^{x_s^k} (1 - D_{OSP}^o(p^k) - D_{OSP}^s(p^k))^{m-x_o^k-x_s^k} \tag{8}
\]

where $x_o^k$ is the amount of customers choosing the online channel in the ordering period $k$; $x_s^k$ is the amount of customers choosing the offline channel in the ordering period $k$; $m!$ denotes the $m$ factorial; $\Psi_{OSP}(x_o^k, x_s^k, p^k)$ denotes the probability that
the amount of customers choosing the online channel is \( x^k_o \) and the amount of customers choosing the offline channel is \( x^k_s \) in the ordering period \( k \) under the OSP policy.

The objective of the dual-channel retailer is to maximize the expected profit under the OSP policy over \( n \) ordering periods. Therefore, the problem can be formulated as below:

\[
\max_{p^k, q^k} \sum_{k=1}^{n} E \left[ (p^k - p_s) x^k_o + (p^k + r) x^k_s - c(I^k + q^k - x^k_o - x^k_s)^+ \right]
\]

\[
- w(-I^k - q^k + x^k_o + x^k_s)^+ \right]
\]

\[
\max_{p^k, q^k} \sum_{k=1}^{n} \sum_{x_o^k=0}^{m-x_s^k} \sum_{x_s^k=0}^{m-x_o^k} \left[ (p^k - p_s) x^k_o + (p^k + r) x^k_s - c(I^k + q^k - x^k_o - x^k_s)^+ \right]
\]

\[
- w(-I^k - q^k + x^k_o + x^k_s)^+ \right] \Psi^{OSP}(x^k_o, x^k_s, p^k)
\]

\[
s.t. \quad I^k = I^{k-1} + q^{k-1} - x^{k-1}_o - x^{k-1}_s 
\]

\[
- n \cdot m \leq I^k \leq n \cdot m 
\]

\[
\max(c, p_s) < p^k < v
\]

where \( q^k \) is the ordering quantity of products in the ordering period \( k \); \( I^k \) is the inventory level of products at the beginning of the ordering period \( k \); \( w \) is the unit shortage cost; \( c \) is the unit inventory holding cost; \( (x)^+ \) denotes \( \max(x, 0) \).

In Eq. (9a), the first two terms are the sale revenues from the online and offline channels under the OSP policy, respectively; the third term is the inventory holding costs under the OSP policy; and the last term is the shortage costs of products under the OSP policy.

### 4.2. The pricing and ordering problem under the WSP policy

In this situation, the customers can also choose online channel and offline channel to purchase products at the selling price \( p^k \). However, the online customers need to afford the shipping price \( p_s \). Therefore, we obtain the utility functions of an individual customer under the WSP policy: \( u_o = v - p^k - h_o \) (i.e., offline utility function under the WSP policy) and \( u_{WSP} = v - p^k - p_s - h_o \) (i.e., online utility function under the WSP policy). Similar to the WSP policy, we assume that the purchasing cost of the customers from the two channels (i.e., \( h_o \) and \( h_s \)) are in the set \([0, H] \times [0, H]\), where \( H > v - p^k \geq 0 \).

Based on the principle of maximum utility, a customer has three choices depending on different online purchasing cost \( h_o \) and offline purchasing cost \( h_s \), as in shown Figure 3. From Figure 3, the probability that a customer chooses the online channel under the WSP policy is the ratio of the area of “online” to the total area (i.e., \( H \times H \)). Similarly, the probability that a customer chooses the offline channel under the WSP policy is the ratio of the area of “offline” to the total area (i.e., \( H \times H \)). Further, we can obtain the probabilities of a customer’s choices for the online and offline channels under the WSP policy as below:

\[
D^{WSP}_o(p^k) = \frac{1}{2H^2} \left[ (2H - v + p^k)(v - p^k - p_s) + 2Hp_s \right]
\]

\[
D^{WSP}_o(p^k) = \frac{1}{2H^2} (2H - v + p^k)(v - p^k - p_s)
\]
where $D_{s}^{WSP}$ is the probability that a customer chooses the offline channel and $D_{o}^{WSP}$ is the probability that a customer chooses the online channel under the WSP policy. Compared with the OSP policy, we can find that some online customers may be transferred to the offline channel or leave when customer needs to afford the shipping cost. Furthermore, from Eqs. (10) and (11), we can obtain Proposition 2:

**Proposition 2.** Under the WSP policy, the probabilities of the customer’s choice satisfy:

1. The probability of the customer’s choice for the online channel (i.e., $D_{o}^{WSP}$) is less than the probability of choosing the offline channel (i.e., $D_{s}^{WSP}$).
2. With the increase of the shipping price, the probability of the customer’s choice for the online channel (i.e., $D_{o}^{WSP}$) decreases while the probability of the customer’s choice for the offline channel (i.e., $D_{s}^{WSP}$) increases.

**Proof.** From Figure 3, the area of “online” is always smaller than the area of “offline”, which leads to $D_{o}^{WSP} < D_{s}^{WSP}$. Further, the first order partial derivatives of those probabilities with respect to $p_s$ are

\[
\frac{\partial D_{o}^{WSP}}{\partial p_s} = \frac{1}{2H^2} (v - p^k) > 0 \tag{12}
\]

\[
\frac{\partial D_{s}^{WSP}}{\partial p_s} = \frac{1}{H^2} (p_s - H) < 0 \tag{13}
\]

Eqs. (12) and (13) prove the third conclusion of this proposition. \qed
Further, based on Eqs. (10) and (11), the probability of demand of the online and offline channels under the WSP policy can be obtained as below:

$$
\Psi_{WSP}^{k}(x_o^k, s^k, p^k) = \frac{m!}{x_o^k!(m - x_o^k)!} \left(D_o WSP(p^k)x_o^k\right) \left(D_s WSP(p^k)x^k - D_s WSP(p^k) m - x_o^k - x^k\right)
$$

where $\Psi_{WSP}^{k}(x_o^k, s^k, p^k)$ denotes the probability that the amount of customers choosing the online channel is $x_o^k$ and the amount of customers choosing the offline channel is $x^k$ in the ordering period $k$ under the WSP policy. Then, we can present the optimal pricing and ordering problem under the WSP policy over $n$ ordering periods that maximizes the expected of profit.

$$
\max_{p^k, q^k} \sum_{k=1}^{n} \mathbb{E}\left[p^k x_o^k + (p^k + r) x^k - c(I^k + q^k - x_o^k - x^k)^+\right] - w(-I^k - q^k + x_o^k + x^k)^+
$$

$$
= \max_{p^k, q^k} \sum_{k=1}^{n} \sum_{x_o^k=0}^{m} \sum_{x^k=0}^{m-x_o^k} \left[p^k x_o^k + (p^k + r) x^k - c(I^k + q^k - x_o^k - x^k)^+\right] - w(-I^k - q^k + x_o^k + x^k)^+ \Psi_{WSP}^{k}(x_o^k, s^k, p^k)
$$

s.t.  \ \ I^k = I^{k-1} + q^{k-1} - x_o^{k-1} - x^k\quad (15b)

$$
- n \cdot m \leq I^k \leq n \cdot m
$$

$$
\text{max}(c, p_s) < p^k < v - p_s
$$

In the objective function Eq. (15a), the first two terms are the sale revenues from the online and offline channels under the WSP policy, respectively; the third term is the inventory holding costs under the WSP policy; and the last term is the shortage costs of products under the WSP policy.

4.3. The pricing and ordering problem under the CFP policy. In practice, the retailers such as Suning.com and Fresh Hema have provided the purchase-more-and-avoid-shipping-price policy for online channel to increase the shopping desire of customer and stimulate the demand of product. In this research, we consider a simple policy that purchasing two products with free shipping but purchasing one product with shipping cost. This is a conditional free shipping (CFP) policy. Under the CFP policy, the customer has four choices to purchase products: (1) the customer purchases products through the offline channel; (2) the customer purchases one product through the online channel and needs to afford the shipping price; (3) the customer purchases two products through the online channel to avoid affording the shipping price; (4) the customer does not purchase any products.

Therefore, we obtain the utility functions of an individual customer under the CFP policy: $u_s = v - p^k - h_s$ (i.e., offline utility function under the CFP policy), $u_{WSP} = v - p^k - p_s - h_o$ (i.e., online utility function for purchasing one product under the CFP policy) and $u_{CFP} = (1 + \theta) v/2 - p^k - h_o/2$ (i.e., online utility function for purchasing two products under the CFP policy). Similar to the WSP policy, we assume that the purchasing cost of the customers from the two channels (i.e., $h_o$ and $h_s$) are in the set $[0, H] \times [0, H]$, where $H > v - p^k - p_s \geq 0$. To ensure the customers consider the CFP policy, we assume that $H > (1 + \theta) v - 2p^k \geq 0, \theta \in [0, 1]$. Based
on different values of discount parameters of the customer’s utility for the second product (i.e., $\theta$), there are several cases in this situation.

4.3.1. $0 < \theta < (p^k - p_s)/v$. In this case, because of less utility of customer for the second product, the customers do not consider purchasing the second product to free shipping price. Therefore, a customer has three choices depending on different online purchasing cost $h_o$ and the offline purchasing cost $h_s$, as shown in Figure 4. From Figure 4, the probability that a customer chooses the online channel is the ratio of the area of “online” to the total area (i.e., $H \times H$). Similarly, the probability that a customer chooses the offline channel is the ratio of the area of “offline” to the total area (i.e., $H \times H$). Therefore, the probabilities of a customer’s choices for the online and offline channels in this case are as bellow:

$$D_{s1}^{CFP}(p^k) = \frac{1}{2H^2} \left[ (2H - v + p^k)(v - p^k - p_s) + 2Hp_s \right]$$  \hspace{1cm} (16)  

$$D_{o1}^{CFP}(p^k) = \frac{1}{2H^2}(2H - v + p^k - p_s)(v - p^k - p_s)$$  \hspace{1cm} (17)  

where $D_{s1}^{CFP}$ denotes the probability that a customer chooses the offline channel; $D_{o1}^{CFP}$ denotes the probability that a customer chooses for the online channel. We can find that this case is similar to the situation under the WSP policy. Therefore, the probabilities of the customer’s choice under this case also satisfy Proposition 2.

4.3.2. $(p^k - p_s)/v < \theta < (v - 2p_s)/v$. In this case, the online customer will consider the policy of purchasing one product and paying the shipping price, or another policy of purchasing two products without paying the shipping cost. Therefore, a customer has four choices depending on different online purchasing cost $h_o$ and offline purchasing cost $h_s$, as shown in Figure 5. From Figure 5, the probability of a customer’s choice is the ratio of the corresponding choice’s area to the total area.
(i.e., $H \times H$). Therefore, we can obtain the probabilities of a customer’s choices for the online and offline channels in this case as below:

$$D_{CFP}^{s2} (p^k) = \frac{1}{H^2} \left[ \left(1 - \theta\right)v \left(1 - \theta\right)v - 2p_s + (\theta v - p^k + p_n)(2v - \theta v - p_s - p^k) + (v - p^k)(H - v - \theta v + 2p^k) \right]$$

$$D_{CFP}^{o2} = \frac{1}{2H^2} (2H - (1 - \theta)v)((1 - \theta)v - 2p_s)$$

$$D_{CFP}^{o2f2} (p^k) = \frac{1}{H^2} \left( 2H + p^k + p_n - (2 - \theta)v \right) \left( \theta v - p^k + p_s \right)$$

where $D_{CFP}^{s2}$ denotes the probability that a customer chooses for the offline channel; $D_{CFP}^{o2}$ denotes the probability that the customer buys one product through the online channel; $D_{CFP}^{o2f2}$ denotes the probability that the customer buys two products through the online channel. From Eqs. (18)-(20), we can obtain Proposition 3:

**Proposition 3.** Under the CFP policy, if $\left( p^k - p_s \right)/v < \theta < \left( v - 2p_s \right)/v$ the probabilities of the customer’s choice satisfy:

1. With the increase of the price of product, the probabilities of the customer’s choice (i.e., $D_{CFP}^{s2}$ and $D_{CFP}^{o2f2}$) decrease.

2. With the increase of the shipping price, the probability of the customer’s choice for the offline channel (i.e., $D_{CFP}^{o2}$) and the probability of the customer’s choice for buying two products through online channel (i.e., $D_{CFP}^{o2f2}$) increase.

**Proof.** Employing Eqs. (18)-(20), the first order partial derivatives of those probabilities with respect to $p^k$ are

$$\frac{\partial D_{CFP}^{s2}}{\partial p^k} = \frac{1}{H^2} \left( v + \theta v - 2p^k - H \right)$$

$$\frac{\partial D_{CFP}^{o2f2}}{\partial p^k} = \frac{2v - 2p^k - 2H}{H^2}$$

Since $H > (1 + \theta)v - 2p^k \geq 0$, those first order partial derivatives are less than zero, which proves the first conclusion of this proposition. Further, the first order partial derivatives of those probabilities with respect to $p_s$ are

$$\frac{\partial D_{CFP}^{s2}}{\partial p_s} = \frac{1}{H^2} \left( (1 - \theta)v - 2p_s \right)$$

$$\frac{\partial D_{CFP}^{o2f2}}{\partial p_s} = \frac{2(H + p_n - (1 - \theta)v)}{H^2}$$

In this case, $\left( p^k - p_s \right)/v < \theta < \left( v - 2p_s \right)/v$. Therefore, the first order partial derivatives of $D_{CFP}^{s2}$ and $D_{CFP}^{o2f2}$ are greater than zero, which proves the second conclusion of this proposition.

From Proposition 3, when the shipping price is higher, the customers tend to purchase more products to avoid the shipping price. Therefore, the CFP policy will stimulate the demand of products for the dual-channel retailing. Besides, although $D_{CFP}^{o2}$ does not include the price of product (i.e., $p^k$) directly, $D_{CFP}^{o2}$ is influenced by the price of product (i.e., $p^k$), since the price of product (i.e., $p^k$) determines the range of discount parameter $\theta$ in this case.
4.3.3. \((v - 2p_s)/v < \theta < (v - 2p_s)/v\). When customers have the higher utility for the second product (i.e., the value of \(\theta\) is greater), the customers do not consider purchase only one product. Therefore, in this case, a customer has three choices depending on different online purchasing cost \(h_o\) and offline purchasing cost \(h_s\), as shown in Figure 6. From Figure 6, the probability of a customer’s choices is the ratio of the corresponding choices’ area to the total area (i.e., \(H \times H\)). Further, we can obtain the probabilities of a customer’s choices for the online and offline channels in this case as follows:

\[
D_{CFP}^{online}(p^k) = \frac{1}{H^2} \left[ \frac{(2H + 2p^k - v - \theta v)(v + \theta v - 2p^k)}{4} + \frac{(1 - \theta)vH}{2} \right] \quad (25)
\]

\[
D_{CFP}^{offline}(p^k) = \frac{(v + \theta v - 2p^k)}{2H^2} \left( 2H + p^k - \frac{(3 - \theta)v}{2} \right) \quad (26)
\]

where \(D_{CFP}^{online}\) denotes the probability that the customer chooses the offline channel, and \(D_{CFP}^{offline}\) denotes the probability that the customer buys two products through the online channel. Although the probabilities of the customer’s choices (i.e., \(D_{CFP}^{online}\) and \(D_{CFP}^{offline}\)) do not include the shipping price directly, the shipping price determines the range of discount parameter \(\theta\) in this case, which affects the probabilities of the customer’s choice. From Eqs. (25) and (26), we can obtain Proposition 4.

**Proposition 4.** Under the CFP policy, if \((v - 2p_s)/v < \theta < (v - 2p_s)/v\), the probabilities of the customer’s choice satisfy:

1. The probability of purchasing only one product through the online channel is zero.
2. With the increase of the price of product, the probabilities of the customer’s choice (i.e., \(D_{CFP}^{online}\) and \(D_{CFP}^{offline}\)) decrease.
Proof. From Figure 6, we can easily obtain the first conclusion of this proposition. Employing Eqs. (25) and (26), the first order partial derivatives of those probabilities with respect to $p^k$ are

$$\frac{\partial D_{s3}^{CFP}}{\partial p^k} = \frac{2}{H^2} (v + \theta v - 2p^k - H) < 0$$

(27)

$$\frac{\partial D_{of3}^{CFP}}{\partial p^k} = \frac{2(v - p^k - H)}{H^2} < 0$$

(28)

From Eqs. (27) and (28), we can easily obtain Proposition 4(2).

---

**Figure 6.** The customer’s choices for purchasing channels for $(v - 2p_s)/v < \theta < 1$

Based on the above analysis, different values of discount parameter of the customer’s utility for the second product (i.e., $\theta$) leads to different cases under the CFP policy. In general, different customers have different values of discount parameter $\theta$. Therefore, we assume that discount parameter of utility for the second product (i.e., $\theta$) is uniformly distributed in [0, 1]. Further, we can obtain the probabilities of a customer’s choice for the online and offline channels under the CFP policy as below.

$$D_s^{CFP}(p^k) = \int_0^{\frac{v - p^k}{v}} D_{s1}^{CFP}(p^k) d\theta + \int_{\frac{v - p^k}{v}}^{\frac{2p_s}{v}} D_{s2}^{CFP}(p^k) d\theta + \int_{\frac{2p_s}{v}}^{1} D_{s3}^{CFP}(p^k) d\theta$$

$$= (p^k)^3 - (3p_s + 6v)(p^k)^2 + (3p_s^2 + 9v^2 - 6Hv)p^k$$

$$+ \frac{3p_s^3 - 6p_s^2v + 3p_s v^2 - 4v^3 + 6Hv^2}{6vH^2}$$

(29)
\[ D^{\text{CFP}}_o (p^k) = \int_0^{v-p_s} D^{\text{CFP}}_{o1} (p^k) d\theta + \int_{v-p_s}^{v-2p_s} D^{\text{CFP}}_{o2} (p^k) d\theta \]
\[ = (v - p^k - p_s) \left[ 2(p_k)^2 + 5p_s^2 + 3H(p_k - 9Hp_s - 5p_k p_s + v(3H - p^k + 2p_s - v)) \right] \]
\[ \frac{6vH^2}{6vH^2} \]

\[ D^{\text{CFP}}_o (p^k) = \int_{v-2p_s}^{2v} D^{\text{CFP}}_{o3} (p^k) d\theta + \int_{v-2p_s}^{v-2p_s} D^{\text{CFP}}_{o4} (p^k) d\theta \]
\[ = \frac{1}{2H^2} \left[ 8p_s (H + p^k) - 4p_s v + \frac{2(p_k + p_s - v)^2(3H + 2p_k - p_s - 2v)}{3v} \right] \]
\[ - \frac{4p_s (3(p_k)^2 + 6Hp^k - p_s^2 + 3Hp_s)}{3v} \]

(30)

where \( D^{\text{CFP}}_s \) denotes the probability that the customer chooses the offline channel under the CFP policy; \( D^{\text{CFP}}_o \) denotes the probability that the customer buys one product through the online channel under the CFP policy; \( D^{\text{CFP}}_{o2} \) denotes the probability that the customer buys two products through the online channel under the CFP policy. Further, based on the Eqs. (29)-(31), the probability of demand of the online and offline channels under the CFP policy can be obtained as below:

\[ \Psi^{\text{CFP}}(x^k_{o}, x^k_{of}, x^k_s, p^k) = \frac{m!}{x^k_{o}!x^k_{of}!x^k_s!} \frac{1}{(m - x^k_{o} - x^k_s - x^k_{of})!} \left( D^{\text{CFP}}_o (p^k) \right)^x^k_{o} \times \left( D^{\text{CFP}}_o (p^k) \right)^x^k_{of} \times \left( D^{\text{CFP}}_o (p^k) \right)^x^k_{s} \times \left( 1 - D^{\text{CFP}}_o (p^k) - D^{\text{CFP}}_o (p^k) - D^{\text{CFP}}_o (p^k) \right)^{m - x^k_{o} - x^k_{of} - x^k_{s}} \]

(32)

where \( \Psi^{\text{CFP}}(x^k_{o}, x^k_{of}, x^k_s, p^k) \) denotes the probability that the amount of customers choosing that buying one product through online channel is \( x^k_{o} \), the amount of customers choosing that buying two product through online channel is \( x^k_{of} \) and the amount of customers choosing the offline channel is \( x^k_s \) in the ordering period \( k \) under the CFP policy. Then, we can present the optimal pricing and ordering problem under the CFP policy, whose objective is to maximize the expected of profit.

\[ \max_{p^k, q^k} \sum_{k=1}^{n} E \left[ p^k x^k_{o} + 2 \left( p^k - p_s \right) x^k_{of} + (p^k + r) x^k_s - c \left( I^k + q^k - x^k_{o} - 2x^k_{of} - x^k_{s} \right)^+ \right] \]
\[ - w \left( - I^k - q^k + x^k_{o} + x^k_{of} + 2x^k_{s} \right)^+ \right] \]
\[ = \max_{p^k, q^k} \sum_{k=1}^{n} \sum_{m=0}^{m} \sum_{m-k-x^k_{of}=0}^{m-k-x^k_{of}} \sum_{x^k_{s}=0}^{x^k_{s}} \left[ p^k x^k_{o} + 2 \left( p^k - p_s \right) x^k_{of} + (p^k + r) x^k_s \right] \]
\[ - c \left( I^k + q^k - x^k_{o} - 2x^k_{of} - x^k_{s} \right)^+ \]
\[ - w \left( - I^k - q^k + x^k_{o} + x^k_{of} + 2x^k_{s} \right)^+ \] \( \Psi^{\text{CFP}}(x^k_{o}, x^k_{of}, x^k_s, p^k) \)

(33a)
Optimal Pricing and Ordering Strategies

\[ s.t. \quad I^k = I^{k-1} + q^{k-1} - x_o^{k-1} - 2x_{of}^{k-1} - x_s^{k-1} \] (33b)
\[ -2n \cdot m \leq I^k \leq 2n \cdot m \] (33c)
\[ \max(c, p_s) < p^k < v - p_s \] (33d)

In Eq. (33a), the first three terms are the sale revenues from the online and offline channels under the CFP policy, respectively; the fourth term is the inventory holding costs under the CFP policy, respectively; and the last term is the shortage costs of products under the CFP policy.

5. Bellman equation and DDPG method for the pricing and ordering problem under the different shipping policies. In this section, based on the presented pricing and ordering problems under the different shipping policies in Section 4, we get the Bellman equations for the three different shipping policies, respectively. Further, we employ the deep deterministic policy gradient method (i.e., DDPG) combining the deep Q-learning method and the actor-critic based policy gradient algorithm to solve those Bellman equations to obtain the optimal pricing and ordering policies under the three different shipping policies.

5.1. Bellman equation under different shipping policies.

5.1.1. Bellman equation under the OSP policy. Under the OSP policy, we first define the reward function as follows:
\[ R_{OSP} [I^k, u(I^k)] = (p^k - p_s) x_o^k + (p^k + r) x_s^k - c(I^k + q^k - x_o^k - x_s^k)^+ - w(-I^k - q^k + x_o^k + x_s^k)^+ \] (34)

where \( u(I^k) \) is the decision variables denoting the ordering quantity \( q^k \) and the price of product \( p^k \) under the state \( I^k \) (i.e., \( u(I^k) = (p^k, q^k) \)). Therefore, the pricing and ordering problem under the OSP policy (i.e., Eq. (9)) can be rewritten as

\[ \max_u \sum_{k=1}^{n} E \{ R_{OSP} [I^k, u(I^k)] \} \] (35)

Further, in order to solve the problem Eq. (35) and obtain the optimal pricing and ordering policy under OSP policy, the maximum expected total profit from ordering period \( k \) to ordering period \( n \) can be described as follows:

\[ J_{OSP}^k (I^k) = \max_u \sum_{j=k}^{n} E \{ R_{OSP} [I^j, u(I^j)] \} \] (36)

Then combining Eq. (34) with Eq. (36), the pricing and ordering problem under the OSP policy (i.e., Eq. (35)) can be formulated as the following Bellman equation:

\[ J_{OSP}^k (I^k) = \max_{u(I^k)} \{ R_{OSP} [I^k, u(I^k)] + J_{OSP}^{k+1} (I^{k+1}) \} , \quad k = 1, \cdots, n-1 \] (37)

In Subsection 5.2, employing the DDPG strategy to solve the Bellman equation (i.e., (37)), we can obtain the optimal pricing and ordering policy under the OSP policy.
5.1.2. Bellman equation under the WSP policy. Similar to the OSP policy, the reward function under the WSP policy can be defined as follows:

\[
R^{WSP}[I^k, u(I^k)] = p^k x_o^k + (p^k + r) x_s^k - c(I^k + q^k - x_o^k - x_s^k)^+ - w(-I^k - q^k + x_o^k + x_s^k)^+ \tag{38}
\]

Therefore, the pricing and ordering problem under the WSP policy (i.e., Eq. (15)) can be rewritten as below

\[
\max_u \sum_{k=1}^n E \{ R^{WSP}[I^k, u(I^k)] \} \quad \text{s.t.} \quad \begin{align*}
I^k &= I^{k-1} + q^{k-1} - x_o^{k-1} - x_s^{k-1} \\
-n \cdot m &\leq I^k \leq n \cdot m \\
\max(c, p_s) &< p^k < v
\end{align*} \tag{39}
\]

For solving the problem Eq. (39) and obtaining the optimal pricing and ordering policy under WSP policy, the maximum expected total profit from ordering period \(k\) to ordering period \(n\) can be described as follows:

\[
J^{WSP}_k(I^k) = \max_u \sum_{j=k}^n E \{ R^{WSP}[I^j, u(I^j)] \} + J^{WSP}_{k+1}(I^{k+1}) \tag{40}
\]

Then, by combining Eq. (38) with Eq. (40), the pricing and ordering problem under the WSP policy (i.e., Eq. (39)) can be formulated as the following Bellman equation:

\[
J^{WSP}_k(I^k) = \max_{u(I^k)} \sum_{j=k}^n E \{ R^{WSP}[I^j, u(I^j)] + J^{WSP}_{k+1}(I^{k+1}) \}, \quad k = 1, \cdots, n-1 \tag{41}
\]

5.1.3. Bellman equation under the CFP policy. Different from the OSP policy and the WSP policy, the reward function under the CFP policy can be defined as follows:

\[
R^{CFP}[I^k, u(I^k)] = p^k x_o^k + 2(p^k - p_s) x_{of}^k + (p^k + r) x_s^k - c(I^k + q^k - x_o^k - 2 x_{of}^k - x_s^k)^+ \\
-w(-I^k - q^k + x_o^k + 2 x_{of}^k + x_s^k)^+ \tag{42}
\]

Under the CFP policy, the online customer can purchase two products for free shipping, or purchase one product and afford the shipping price. Further, we can rewritten the pricing and ordering problem under the CFP policy (i.e., Eq. (33)) as below

\[
\max_u \sum_{k=1}^n E \{ R^{CFP}[I^k, u(I^k)] \} \quad \text{s.t.} \quad \begin{align*}
I^k &= I^{k-1} + q^{k-1} - x_o^{k-1} - 2 x_{of}^{k-1} - x_s^{k-1} \\
-2n \cdot m &\leq I^k \leq 2n \cdot m \\
\max(c, p_s) &< p^k < v - p_s
\end{align*} \tag{43}
\]
Then, we define the maximum expected total profit from ordering period $k$ to ordering period $n$ as follows:

$$J_{CFP}^k(I_k) = \max_{u(I_k)} \sum_{j=k}^{n} E \left\{ R_{CFP}^{j} \left[ I_j, u(I_j) \right] \right\}$$ (44)

Therefore, by combining Eq. (42) with Eq. (44), the pricing and ordering problem under the CFP policy (i.e., Eq.(43)) can be formulated as the following Bellman equation:

$$J_{CFP}^k(I_k) = \max_{u(I_k)} E \left\{ R_{CFP}^{k} \left[ I_k, u(I_k) \right] + J_{CFP}^{k+1}(I_{k+1}) \right\}, \quad k = 1, \cdots, n-1$$ (45)

Similar to the OSP policy and the WSP policy, we can also employ the DDPG strategy to solve the Bellman equation (i.e., Eq (45)) to obtain the optimal pricing and ordering policy under the CFP policy.

5.2. **DDPG method under different shipping policies.** Before solving those Bellman equations in subsection 5.1, we first define the action-value function for different shipping policies as follows:

$$Q_i^k(I_k, u(I_k)) = E \left\{ R_i^{k} \left[ I_k, u(I_k) \right] + J_i^{k+1}(I_{k+1}) \right\}$$ (46)

In Eq.(46), when the subscript $i$ is OSP, WSP or CFP, this equation denotes the action-value function under the OSP policy, the WSP policy or the CFP policy respectively.

Therefore, the maximum expected total profit from the ordering period $k$ to the ordering period $n$ under different shipping policies (i.e., $J^i(I_k)$) can be rewritten by using the action-value function as follows:

$$J_k^i(I_k) = \max_{u(I_k)} Q_k^i(I_k, u(I_k))$$ (47)

Then, the obtained Bellman equations in subsection 5.1 can be rewritten in terms of the action-value function as below:

$$Q_k^i(I_k, u(I_k)) = E \left\{ R^i \left[ I_k, u(I_k) \right] + \gamma \max_{u(I_{k+1})} Q_k^i(I_{k+1}, u(I_{k+1})) \right\}$$ (48)

where $\gamma$ is the discount factor.

For the Bellman equation (i.e., Eq. (48)), we can employ the approximate dynamic method (i.e., deep $Q$-learning method) to solve it. However, we can find that since the deep $Q$-learning method needs to go through all actions to gain the maximize action-value function at the next state, it can be adopted to solve the problem with discrete action space effectively. However, when the action space is continuous and there are more decision variables, this method will lead to the curse of dimensionality and spend more time. Notably, we have two decision variables (i.e., $q^k$ and $p^k$), and the action space is continuous in this paper. Therefore, we employ the deep deterministic policy gradient method (i.e., DDPG) combining deep $Q$-learning method and the actor-critic-based policy gradient algorithm [21] to solve those problems.

This method has some features: (1) employing the actor network to the approximate action-value function; (2) employing the critic network to train the deterministic policy by choosing the action under the state; (3) using experience replay and fixed target network to increase the speed of the converge performance and reduce time consumption [21]. Therefore, there are four neural networks in this method.
Algorithm 1 DDPG for the pricing and ordering problem

**Initialization:**
- Initialize the index of iteration \( l \) and the index of ordering period \( k \) to zero;
- Construct feedforward neural critic network and feedforward neural actor network with random parameters \( \theta^Q \) and \( \theta^\mu \);
- Initialize target critic network and target actor network with \( \bar{\theta}^Q \leftarrow \theta^Q \) and \( \bar{\theta}^\mu \leftarrow \theta^\mu \);
- Initialize the experience playback memory \( D \) to size \( W \);
- While (iteration \( l < \text{Max iteration} \))
  - Initialize the state \( I^0 \) and the action exploration process \( N \);
  - For \( k = 1, n \) do
    - Based on the state \( I^k \) to select action \( u(I^k) = \mu(I^k|\theta^\mu) + N^k \)
    - Based on the probability of demand (i.e., Eqs.(8), (14) and (32)) and the selected action, obtaining the random demand of product;
    - Based on the Eq. (34) or Eq. (38) or Eq. (42), computing the reward function \( R^i \);
    - Storing \( [I^k, u(I^k), R^i(I^k, u(I^k))], I^{k+1} \] to experience playback memory \( D \);
    - Randomly selecting a subset with fixed size \( G \) from memory \( D \);
    - Based on the selected \( [I_j, u(I_j), R^i(I_j, u(I_j))], I_{j+1} \], update the critic network by minimizing the loss function:
      \[
      L = \frac{1}{G} \sum_j \left( Q^i(I_j, u(I_j)|\theta^Q) - y^j \right)^2
      \]
      \( y^j = R^i[I_j, u(I_j)] + \gamma \bar{Q}^i(I_{j+1}, \bar{\mu}(I_{j+1}|\theta^\bar{\mu})|\theta^\bar{Q}) \)
    - Update the actor network by using the policy gradient method:
      \[
      \nabla_{\theta^\mu} J = \frac{1}{G} \sum_j \nabla_{\theta^\mu} Q^i(I, a|\theta^Q) \bigg|_{I=I_j, a=\mu(I_j)} \nabla_{\theta^\mu} \mu(I|\theta^\mu) \bigg|_{I=I_j}
      \]
    - Employ the Eq. (52) to update the target network;
  - end for
- end while

(i.e., the online actor network, the online critic network, the target actor network and the target critic network).

In the critic network, the action-value function is updated by the approximate function \( Q^i(I^k, u(I^k)|\theta^Q) \), which can be optimized by minimizing the following loss function
\[
L(\theta^Q) = E[Q^i(I^k, u(I^k)|\theta^Q) - y^k]^2
\]  \hspace{1cm} (49)
where
\[
y^k = R^i[I^k, u(I^k)] + \gamma \bar{Q}^i(I^{k+1}, \bar{\mu}(I^{k+1}|\theta^\bar{\mu})|\theta^\bar{Q})
\]  \hspace{1cm} (50)
\( \theta^Q \) and \( \theta^\bar{Q} \) denote the network parameters of actor network \( Q^i \) and target actor network \( \bar{Q}^i \) respectively; \( \theta^\bar{\mu} \) denotes the network parameter of target critic network \( \bar{\mu} \).

The actor network employs the policy gradient method to update the policy choosing action based on the critic network. The policy’s updating gradient is the
following equation:

\[
\nabla_{\theta^\mu} J = E \left[ \nabla_{\theta^\mu} Q^i \left( I, a \mid \theta^Q \right) \vert I = I^k, a = \mu(I_k) \theta^\epsilon \right] \\
= E \left[ \left( \nabla_a Q^i \left( I, a \mid \theta^Q \right) \vert I = I^k, a = \mu(I_k) \right)^T \nabla_{\theta^\mu} u \left( I \mid \theta^\mu \right) \vert I = I^k \right] 
\]

(51)

where \( \theta^\mu \) denotes the network parameter of critic network \( \mu \). Besides, target nets are updated with constant \( \tau \).

\[
\theta^Q \leftarrow \tau \theta^Q + (1 - \tau) \bar{\theta}^Q \\
\theta^\mu \leftarrow \tau \theta^\mu + (1 - \tau) \bar{\theta}^\mu
\]

(52)

Further, we present the details of the DDPG strategy for the pricing and ordering problems under different shipping policies can be found in Table 1.

6. Numerical experiments. In this section, we employ the proposed method in Section 5.2 (i.e., DDPG method) to obtain the optimal pricing and ordering policies for the different shipping policies (i.e., OSP, WSP and CFP). Further, we conduct comparison study for three different shipping policies. Finally, based on the analysis, we obtain some managerial insights.

The basic parameters are \( v = 1.6, H = 2.2, m = 100, n = 5, w = 0.1, c = 0.2, r = 0.1 \) and \( p_e = 0.2 \). The optimal price, ordering quantity and total profit under the OSP policy for different shipping prices are shown in Table 2. In Table 2, \( \bar{p} \) and \( \bar{q} \) denote the average price and the average ordering quantity during \( n \) ordering periods.

Table 2 shows that during \( n \) ordering periods, the price or ordering quantity does not fluctuate very much. With the increase of shipping price, the optimal price increases, while the ordering quantity and total profit decrease. The reason is that when the shipping price increases under the OSP policy, the retailer needs to afford the higher shipping price which leads to the increase of the price of product to obtain more profit. However, under the OSP policy, the higher price incurs the retailer to afford more cost, which leads to the decrease of the total profit. From Proposition 1, higher price of product leads to the decrease of the probabilities of a customer’s choices for the online and offline channels under the OSP policy. Therefore, the ordering quantity of product will decrease.

For the same dual-channel inventory system, the optimal price, ordering quantity and total profit under the WSP policy for different shipping prices are shown in Table 3. From Table 3, we can find that during \( n \) ordering periods, the price or ordering quantity also does not fluctuate very much, and the increase of shipping price leads to the increase of ordering quantity and the decrease of the price of product and the total profit. This is because under the WSP policy, the customer choosing the online channel needs to afford the shipping price, and the retailer needs to reduce the price of product to attract more customers. On the contrary, under the OSP policy, the increase of the shipping price will lead to the decrease of price of product and the increase of the ordering quantity.

The optimal price, ordering quantity and total profit under the CFP policy for different shipping price are shown in Table 4. Similar to the OSP policy, with the increase of the shipping price, the price of product increases and the ordering quantity and total profit decrease. Under the CFP policy, the retailer needs to afford the shipping price for the online customer purchasing two products, while the online customer purchasing one product needs to afford the shipping price. Therefore,
when the shipping price increases, the retailer increases the price of product to attract more customers to purchase two products to free shipping. However, the higher shipping price increases the probabilities of buying no products, which leads to the decrease of ordering quantity.

The optimal price, ordering quantity and total profit under different shipping policies for the different shipping price are as shown Table 5. For the price of products, whether the retailer affords the shipping price has great impact on the optimal price of product. We can find that the price of product under the OSP policy is higher than the prices under other policies, because the retailer needs to afford the shipping price under the OSP policy. For the ordering quantity, we can
Table 4. The price, ordering quantity and total profit under the CFP policy for different shipping price $p_s$.

| $k$ | $k = 1$ | $k = 2$ | $k = 3$ | $k = 4$ | $k = 5$ | $\bar{p}$ or $\bar{q}$ |
|-----|---------|---------|---------|---------|---------|----------------|
| Optimal price $p^k$ |          |         |         |         |         |                |
| $p_s = 0.2$ | 0.873   | 0.857   | 0.867   | 0.873   | 0.869   | 0.868         |
| $p_s = 0.4$ | 0.908   | 0.904   | 0.911   | 0.916   | 0.926   | 0.913         |
| $p_s = 0.6$ | 0.990   | 0.980   | 0.970   | 0.990   | 0.980   | 0.982         |
| Optimal ordering quantity $q^k$ |          |         |         |         |         |                |
| $p_s = 0.2$ | 79      | 71      | 75      | 79      | 79      | 77            |
| $p_s = 0.4$ | 72      | 69      | 66      | 65      | 62      | 67            |
| $p_s = 0.6$ | 61      | 64      | 67      | 70      | 58      | 64            |
| Total profit |          |         |         |         |         |                |
| $p_s = 0.2$ | 52.07   | 110.86  | 176.08  | 241.16  | 292.13  |                |
| $p_s = 0.4$ | 49.31   | 94.28   | 142.13  | 183.62  | 229.91  |                |
| $p_s = 0.6$ | 40.07   | 81.73   | 130.55  | 174.87  | 212.56  |                |

Table 5. The comparison results under the different shipping policies.

|            | $p_s = 0.2$ | $p_s = 0.4$ | $p_s = 0.6$ |
|------------|-------------|-------------|-------------|
| Price difference |            |            |             |
| $\bar{p}^{WSP} - \bar{p}^{OSP}$ | -0.027 | -0.175 | -0.321 |
| $\bar{p}^{CFP} - \bar{p}^{WSP}$ | -0.048 | 0.097 | 0.239 |
| $\bar{p}^{CFP} - \bar{p}^{OSP}$ | -0.075 | -0.078 | -0.082 |
| Ordering quantity difference |            |            |             |
| $\bar{q}^{WSP} - \bar{q}^{OSP}$ | -9 | -3 | 2 |
| $\bar{q}^{CFP} - \bar{q}^{WSP}$ | 33 | 20 | 17 |
| $\bar{q}^{CFP} - \bar{q}^{OSP}$ | 24 | 17 | 19 |
| Total profit difference |            |            |             |
| $WSP - OSP$ | -13.78 | -1.75 | 7.1 |
| $CFP - WSP$ | 77.03 | 35.46 | 20.03 |
| $CFP - OSP$ | 63.25 | 33.89 | 27.13 |

find that the ordering quantity under the CFP policy is much more than ordering quantities under other policies, because the retailer under the CFP policy affords the shipping price for the online customers who purchase more products, which stimulates the demand of product.

From the total profit difference for different shipping policies, the CFP policy is better than the others whatever the shipping price is lower or higher. Therefore, the policy that purchasing more products for free shipping is helpful to stimulate the demand of product and further makes the retailer gain more profit. Besides, the OSP policy is much better than the WSP policy for the lower price, while the WSP policy is much better than the OSP policy for the higher price.
Conclusion. In this paper, we investigate the optimal pricing and ordering problem under three different shipping policies. The customers purchase products through the online channel and the offline channel. It is noted that under different shipping policies, the customers from the online channel avoid the shipping price (i.e., OSP policy), afford the shipping policy (i.e., WSP policy) or purchase more products to avoid the shipping price (i.e., CFP policy). Firstly, we model the probabilities of a customer’s choices for the online and offline channels by employing the utility function of the customer, and further obtain the probability of demand for the online and offline channels under the different shipping policies. Based on the probability of demand under the different shipping policies, we model the pricing and ordering problems. Finally, to avoid the curse of dimensionality, the DDPG method is used to solve those problems to obtain the optimal pricing and ordering policy.

From the theoretical analysis and the numerical experiments, the optimal pricing and ordering policy does not fluctuate dramatically. Comparing with the OSP policy and the WSP policy, the probability of demand for the online and offline channels under the CFP policy is bigger. Under the OSP policy, since the retailers need to afford the shipping price, the optimal price increases with the increase of shipping price, which leads to the decrease of the ordering quantity. On the contrary, under the WSP policy, the customers need to afford the shipping price, therefore the retailers should decrease the optimal price to obtain more profit, which leads to the increase of the ordering quantity. Under the CFP policy, the optimal price of product increases with the increase of shipping price. Especially, since the customer can choose the policy of purchasing two products with free shipping under the CFP policy, higher shipping price can stimulate the customer to purchase more products, which further increases the ordering quantity.

Comparing with the OSP policy and WSP policy, the OSP policy is much better than the WSP policy for the lower shipping price, while the WSP policy is much better than the OSP policy for the higher shipping price. Therefore, among the OSP policy and WSP policy, the retailer should afford the shipping price when the shipping price is lower, while the retailer should let the customer to afford the shipping price when the shipping price is higher. Besides, the total profit and the ordering quantity under the CFP policy are much more than those under other policies whenever the shipping price is lower or higher. As for the CFP policy, it can attract customers to purchase more products, and further increases the demand of product to gain more profit. Therefore, we recommend that the dual-channel retailers should choose the CFP policy as the shipping policy.

In the future, we will relax the assumptions in Section 3 to research the pricing and inventory management problem under the different shipping policies. In addition, the model for the situation of single type of products will be extended to the case of multiple types of products. Furthermore, demand substitution will also be taken into consideration for the multiple product-types case. Besides, we will consider the situation that the dual-channel retailers can provide the buy-online-and-pick-up-in-store channel under different shipping policies, which will reveal more managerial insights.

Acknowledgments. This research is supported by National Natural Science Foundation of China (Grant No. 61673109).
REFERENCES

[1] F. Alawneh and G. Q. Zhang, Dual-channel warehouse and inventory management with stochastic demand, *Transportation Research Part E-Logistics and Transportation Review*, 112 (2018), 84–106.

[2] R. Batari, M. Y. Jaber and C. H. Glock, Pricing and inventory decisions in a dual-channel supply chain with learning and forgetting, *Computers and Industrial Engineering*, 136 (2019), 397–420.

[3] R. Becerril-Arreola, M. M. Leng and M. Parlar, Online retailers’ promotional pricing, free-shipping threshold, and inventory decisions: A simulation-based analysis, *European J. Oper. Res.*, 230 (2013), 272–283.

[4] J. X. Chen, L. Liang, D. Q. Yao and S. N. Sun, Price and quality decisions in dual-channel supply chains, *European J. Oper. Res.*, 259 (2017), 935–948.

[5] W. K. Chiang, D. Chhajed and J. D. Hess, Direct marketing, indirect profits: A strategic analysis of dual-channel supply-chain design, *Management Science*, 49 (2003), 1–20.

[6] L. Ding, X. Liu and Y. Xu, Competitive risk management for online Bahncard problem, *J. Ind. Manag. Optim.*, 6 (2010), 1–14.

[7] R. Batarfi, M. Y. Jaber and C. H. Glock, Pricing and inventory decisions in a dual-channel supply chain with learning and forgetting, *Computers and Industrial Engineering*, 136 (2019), 397–420.

[8] C. Fan, Y. M. Liu, X. H. Yang, X. H. Chen and J. H. Hu, Online and offline cooperation under buy-online, pick-up-in-store: Pricing and inventory decisions, *Management Science*, 52 (2014), 1–20.

[9] G. W. Hua, S. Y. Wang and T. C. E. Cheng, Price and lead time decisions in dual-channel supply chains, *European Journal of Operational Research*, 205 (2010), 113–126.

[10] G. Huang, Q. Ding, C. Dong and Z. Pan, Joint optimization of pricing and inventory control for dual-channel problem under stochastic demand, *Ann. Oper. Res.*, 298 (2021), 307–337.

[11] W. Huang and J. M. Swaminathan, Introduction of second channel: Implications for pricing and profits, *European Journal of Operational Research*, 194 (2009), 258–297.

[12] W. H. Huang and Y. C. Cheng, Threshold free shipping policies for internet shoppers, *Transportation Research Part A*, 82 (2015), 193–203.

[13] R. X. Kong, L. Luo, L. X. Chen and M. F. Keblis, The effects of BOPS implementation under different pricing strategies in omnichannel retailing, *Transportation Research Part E*, 141 (2020), 102014, 1–30.

[14] T. T. Li, X. B. Zhao and J. X. Xie, Inventory management for dual sales channels with inventory-level-dependent demand, *Journal of the Operational Research Society*, 66 (2015), 488–499.

[15] T. P. Lillicrap, J. J. Huant, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver and D. Wierstra, Continuous control with deep reinforcement learning, *Computer Science*, 529 (2016), 484–489, arXiv:1509.02971.

[16] B. Liu, R. Zhang and M. D. Xiao, Joint decision on production and pricing for online dual channel supply chain system, *Appl. Math. Model.*, 34 (2010), 4208–4218.

[17] N. M. Modak and P. Kelle, Managing a dual-channel supply chain under price and delivery-time dependent stochastic demand, *European J. Oper. Res.*, 272 (2019), 147–161.
[24] A. Roy, S. S. Sana and K. Chaudhuri, Joint decision on eoq and pricing strategy of a dual channel of mixed retail and e-tail comprising of single manufacture and retail under stochastic demand, *Computers & Industrial Engineering*, **102** (2016), 423–434.

[25] J. K. Ryan, D. Sun and X. Y. Zhao, Coordinating a supply chain with a manufacturer-owned online channel: A dual channel model under price competition, *IEEE Transactions on Engineering Management*, **60** (2013), 247–259.

[26] X. F. Shao, Free or calculated shipping: Impact of delivery cost on supply chains moving to online retailing, *International Journal of Production Economics*, **191** (2017), 267–277.

[27] W. G. Tang, H. T. Li and K. Cai, Optimising the credit term decisions in a dual-channel supply chain, *International Journal of Production Research*, **59** (2021), 4324–4341.

[28] R. H. Wang, G. G. Nan, L. Chen and M. Q. Li, Channel integration choices and pricing strategies for competing dual-channel retailers, *IEEE Transactions on Engineering Management*, (2020), 1–15.

[29] J. Q. Yang, X. M. Zhang, H. Y. Fu and C. Liu, Inventory competition in a dual-channel supply chain with delivery lead time consideration, *Appl. Math. Model.*, **42** (2017), 675–692.

[30] D. Q. Yao, X. H. Yue, X. Y. Wang and J. J. Liu, The impact of information sharing on a returns policy with the addition of direct channel, *International Journal of Production Economics*, **97** (2005), 196–209.

[31] Y. Zhang, H. F. Zhong, Y. Liu and M. H. Huang, Online ordering strategy for the discrete newsvendor problem with order value-based free-shipping, *J. Ind. Manag. Optim.*, **15** (2019), 1617–1630.

[32] B. Zhou, M. N. Katehakis and Y. Zhao, Managing stochastic inventory systems with free shipping option, *European J. Oper. Res.*, **196** (2009), 186–197.

Received July 2021; revised October 2021; early access January 2022.

E-mail address: 230208165@seu.edu.cn
E-mail address: wangz@seu.edu.cn