1-2024

Managing the personalized order-holding problem in online retailing

Shouchang CHEN
Zhejiang University

Zhenzhen YAN
Nanyang Technological University

Yun Fong LIM
Singapore Management University, yflim@smu.edu.sg

Citation

CHEN, Shouchang; YAN, Zhenzhen; and LIM, Yun Fong. Managing the personalized order-holding problem in online retailing. (2024). *Manufacturing & Service Operations Management*. 26, (1), 47-65.

Available at: https://ink.library.smu.edu.sg/lkcsb_research/7176

Managing The Personalized Order-Holding Problem in Online Retailing

Shouchang Chen
School of Management, Zhejiang University, Hangzhou, 310058, China, chenshouchang@zju.edu.cn

Zhenzhen Yan
School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, yanzz@ntu.edu.sg

Yun Fong Lim
Lee Kong Chian School of Business, Singapore Management University, Singapore 178899, yflim@smu.edu.sg

Feb 03, 2023

Problem definition: A significant percentage of online consumers place consecutive orders within a short duration. To reduce the total order arrangement cost, an online retailer may consolidate consecutive orders from the same consumer. We investigate how long the retailer should hold the consumer's orders before sending them to a third-party logistics provider (3PL) for processing. In this order-holding problem, we optimize the holding time to balance the total order arrangement cost and the potential delay in delivery.

Methodology/results: We model the order-holding problem as a Markov Decision Process. We show that the optimal order-holding decisions follow a threshold-type policy that is straightforward to implement: Hold any pending orders if the holding time is within a threshold, or send them to the 3PL otherwise. Whenever the consumer places a new order, the holding time is reset and the threshold is updated based on a cumulative set of her past consecutive orders in her shopping journey. Using a consumer's sequential decision model, we personalize the threshold by finding its closed-form expression in the consumer's order features. We determine the model's coefficients and evaluate the threshold-type policy using the data of the 2020 MSOM Data Driven Research Challenge. Extensive numerical experiments suggest that the personalized threshold-type policy outperforms two commonly-used benchmarks by having fewer order arrangements or shorter holding times. Furthermore, personalizing the order-holding decisions is significantly more valuable for "enterprise" customers.

Managerial implications: Our research suggests a higher threshold for consumers who are more likely to place consecutive orders within a short duration. The consumers' demographic information has a significant effect on the threshold. Specifically, the threshold is higher for "plus" consumers, female consumers, and consumers in the age group of 16-25. The threshold for tier-1 cities is lower than that for tier-2 to tier-4 cities but higher than that for tier-5 cities.

Key words: Online retailing, order holding, personalized threshold-type policy, sequential decision model

1. Introduction

As technology continues to draw consumers from brick-and-mortar stores to online shops, e-commerce becomes an increasingly common way of shopping. In 2020, global retail e-commerce sales reached $4.28 trillion, representing 18% of the total retail sales worldwide (eMarketer 2021). The revenues of e-retailing are projected to grow to $6.39 trillion, making up 21.8% of the total retail sales worldwide in 2024. In contrast to brick-and-mortar retailing, products sold online are handled in *single-item* or *small-quantity* orders before they reach individual consumers. After an order is placed online, it will be assigned to a warehouse where a fulfillment process begins. The impact of order-fulfillment efficiency on an online retailer's profit margin is significant. For example, in 2018 Amazon spent approximately 30 billion dollars on shipping costs (Richter 2019), while its net income was approximately 10 billion dollars (Macrotrends.net 2019). Achieving high order-fulfillment efficiency is therefore crucial for online retailing.

However, it is challenging to efficiently fulfill customer orders in e-commerce. After an order is placed online, it will go through several stages including order-warehouse assignment, order-picking, packing, and delivery before it reaches the consumer. Each stage affects the overall order-fulfillment efficiency. Different stages have been studied in the literature. Xu et al. (2009), Jasin and Sinha (2015), Andrews et al. (2019), and Lim et al. (2021) optimize order-warehouse assignments to minimize fulfillment costs or maximize rewards. Other papers focus on improving warehouse operations including order-picking and packing (Rouwenhorst et al. 2000, De Koster et al. 2007, Gzara et al. 2020). Some researchers study how to consolidate multiple orders from different consumers into a single shipment or split an order into multiple shipments to reduce the total shipping fee, while meeting delivery deadlines (Acimovic and Graves 2015, Wei et al. 2017).

In contrast to the existing literature, we tackle the order-fulfillment challenge from a different perspective: We investigate **how long should an online retailer hold a consumer's consecutive orders before sending them to a third-party logistics provider (3PL) (or an in-house logistics department) for processing**. Typically, the online retailer pays a fixed transaction fee to transmit a batch of orders to the 3PL in each *order arrangement*. This transaction fee does not include the actual order-picking cost in the warehouse and is generally independent of the number of items in the orders. Thus, consolidating consecutive orders from *the same* consumer can potentially reduce the total order arrangement cost substantially.

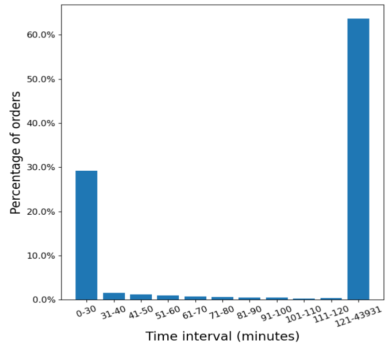


Figure 1 The distribution of the time interval between two subsequent orders

Our study is motivated by Figure 1 based on JD.com's data in *2020 MSOM Data Driven Research Challenge* (Shen et al. 2019). Among orders received by JD.com with a subsequent order placed by the same consumer within one month, Figure 1 shows that around 30% of them have their subsequent order placed within 30 minutes. This is because as product recommendations become increasingly common, consumers with a high search cost are easily tempted to purchase more items after completing an order. Furthermore, memberships such as JD plus and Amazon prime offer free shipping service regardless of the order size. Consumers are free to place multiple orders (instead of one large order) without an extra fee. The tendency of placing consecutive orders motivates us to consolidate orders from the same consumer to reduce the total order arrangement cost. Our research is significant because the JD.com data shows that 1.72% of orders are placed within 30 minutes from the previous order by the same consumer. In 2016, JD.com received 1.6 billion orders (Business Statistics 2021). By consolidating the orders placed within 30 minutes by the same consumer, we can reduce 27.52 (1,600 $\times$ 1.72%) million order arrangements annually, representing a cost reduction of 27.52 million times of the fixed fee per order arrangement.

Another common concern of online retailers is the risk of order-delivery delay. The longer the orders are held, the later they will be delivered, affecting the customer satisfaction level – a key performance indicator. About 43% of consumers choose *next-day delivery*, and 17% of consumers abandon a brand if they wait for too long for their order deliveries (Ed Smith 2020). Online retailers are looking for solutions that not only reduce their total order arrangement cost, but also ensure a high customer satisfaction level. We capture customer satisfaction in our model using *disutility in delivery time*. Since consumers can view their order status on retailers' website, a delay in sending an order to the 3PL will affect customer satisfaction and contribute to the disutility that discourages order-holding. Although consolidating a consumer's consecutive orders can reduce the total order arrangement cost, a long holding time increases the disutility. It is crucial to optimize the holding time to balance the total order arrangement cost and the disutility in delivery time.

In this paper, we study *the order-holding problem*: How long should an online retailer hold a consumer's orders before sending them to a 3PL for processing? Our goal is to balance the total order arrangement cost and the disutility in delivery time. We define the system state in three dimensions: (i) The first represents the features of the consumer's past consecutive orders in her shopping journey (including the consumer's demographic information and the product attributes). (ii) The second is a duration in which no new orders are placed by the consumer since her latest order. (iii) The last indicates whether the consumer holds any pending orders that have not been sent to the 3PL. We model the order-holding problem as a Markov Decision Process (MDP) where the retailer decides in each period whether to continue holding the pending orders (if any) placed by the consumer. Our model incorporates the features of the consumer's past consecutive orders in her shopping journey. Depending on the products ordered by the consumer, the probability of placing another order shortly by the consumer could vary, causing her order-holding decision to change dynamically. For example, a customer who first orders a tape may have some probability to order moving boxes shortly. The retailer may hold the tape order for a longer time to combine it with a potential order of boxes later. However, once the customer also orders the boxes, her probability of placing another order shortly may drop and it may be better for the retailer to send the pending orders to the 3PL immediately.

We summarize the main contributions of our paper as follows:

1. Due to the curse of dimensionality, we do not solve the MDP directly to find the retailer's optimal policy. Instead, we first construct a set of single-dimensional MDPs and show that the optimal policy of the original MDP can be obtained by solving these single-dimensional MDPs. Each single-dimensional MDP is independent of the number of products, which is often huge for online retailing. Therefore, it is more tractable and significantly reduces the computation burden.

2. We show that the retailer's optimal decisions follow a threshold-type policy: *Hold the pending orders if the holding time is within a threshold, or send them to the 3PL otherwise*. If the consumer places a new order, then the holding time is reset and the threshold is updated based on the consumer's cumulative set of past consecutive orders in her shopping journey. We show that the holding threshold depends on the transition probabilities of an associated single-dimensional MDP. We derive a closed-form relation between the transition probabilities and the features of the consumer's past consecutive orders, allowing us to *personalize* each holding threshold. We also show that the optimal personalized threshold is positively correlated with the consumer's probability of placing consecutive orders within a short duration. We derive this probability in terms of the

consumer's order features in her shopping journey. This helps the retailer predict the consumer's probability of placing another order shortly after her latest order.

3. Using the JD.com data, we apply a *maximum likelihood estimation* (MLE) method to find the coefficients of our model. We find that the MLE method accurately predicts the probabilities of our model. We propose a piecewise-linear function that well approximates the optimal threshold. Using the estimated coefficients of this piecewise-linear function, we can quantify the dependence of the threshold on the order features such as the consumer's gender, age group, and location, allowing us to identify some important features. We find that the personalized threshold-type policy outperforms two widely-used benchmarks by having fewer order arrangements or shorter holding times, and optimizing order holding is significantly more valuable for "enterprise" customers.

Section 2 reviews the related literature. Section 3 formulates the order-holding problem. Section 4 applies a consumer's sequential decision model to express the transition probabilities in the order features, and derives the personalized threshold and its piecewise-linear approximation. Section 5 estimates the coefficients of the consumer's sequential decision model. Section 6 processes the JD.com data. Section 7 conducts numerical experiments and discusses insights. Section 8 concludes the paper. All proofs can be found in the online appendix.

2. Literature Review

This paper is related to three streams of literature: (i) order-fulfillment problems, (ii) consumer choice models, and (iii) estimating transition probabilities of Markov chains.

2.1. Order-Fulfillment Problems

Order-fulfillment problems are widely studied in the literature. Xu et al. (2009) consider an order-warehouse assignment problem. By re-evaluating the available inventory at the warehouses and the estimated-to-ship dates, they update the real-time decisions on when and from which warehouse to ship to minimize the number of shipments. Jasin and Sinha (2015) consider a related problem but they optimize decisions in a forward-looking rather than a myopic fashion. Andrews et al. (2019) propose a primal-dual approach to solve an assignment problem between orders and warehouses to maximize the reward. Lim et al. (2021) integrate the replenishment, allocation, and fulfillment decisions for an online retailer in one model. They solve the model using robust optimization approaches. Lei et al. (2018) further incorporate pricing decisions that affect the supply-demand relation when making the order-fulfillment plan.

There is a rich stream of literature on order fulfillment within warehouses. See De Koster et al. (2007) for a comprehensive review. Many strategies have been proposed to improve order-picking, which has the highest operating cost in a warehouse. For example, Gzara et al. (2020) optimize order consolidation in a major e-commerce warehouse based on a comprehensive data analysis. They develop optimal and heuristic approaches to find an effective order-consolidation policy.

Other papers consider shipping assignment. Acimovic and Graves (2015) propose a dynamic program to solve a shipping assignment problem to determine from where to ship items and how multi-item orders can be split into multiple shipments to minimize the average total shipping cost. Wei et al. (2017) study a shipment consolidation problem by addressing a trade-off between the cost reduction from consolidating multiple orders and a potential fee increase due to expedited shipping. In contrast to the above papers, we study how long should an online retailer hold a single consumer's orders before sending them for processing by a 3PL.

Some papers consider consolidating orders from multiple consumers in the fulfillment process (Stenius et al. 2016, Çeven and Gue 2017, Wei et al. 2017, Gzara et al. 2020). These papers assume a stationary order arrival process. In contrast, we consider consolidating orders from *the same* consumer and predict the probability of a coming order by analyzing the consumer's sequential decision process. Comparing with the existing literature on order consolidation, our paper is unique because we incorporate order features to capture different purchase behaviors of different consumers. This leads to a challenging problem on how to model consumer heterogeneity and how to design a personalized consolidation policy. We solve this by applying a consumer sequential choice model to characterize the transition probabilities in a feature-dependent MDP. Different consumers are associated with different sequential choice models and hence different ordering decisions, which leads to a personalized consolidation policy.

2.2. Consumer Choice Model

This paper applies a consumer's sequential decision (choice) model to characterize the transition probabilities of a set of single-dimensional MDPs. Various choice models exist in the literature, including the multinomial logit model (MNL) (Luce 1959), nested multinomial logit model (Williams 1977), Markov chain choice model (Blanchet et al. 2016), exponomial model (Alptekinoğlu and Semple 2016), and representative consumer choice model (Yan et al. 2022). Researchers have extended static choice models to multi-stage choice models (Flores et al. 2019, Gallego et al. 2020, Liu et al. 2020), where a consumer proceeds in multiple stages and the conditional probability at each stage follows an MNL choice probability.

A well-established model for a multi-stage decision process is a Markovian logit choice model (Bell 1995, Akamatsu 1996, 1997), which assumes that the random error terms of the utility model follow independent and identical Gumbel distributions. Ahipaşaoğlu et al. (2019) generalize the model to a distributionally-robust setting assuming only the marginal distributions of the random utility terms are given. Yan (2020) follows this idea to generalize the sequential choice model to characterize multi-item choice behavior. We build our sequential choice model based on Yan (2020) and characterize the relation between a consumer's utility model and the transition probabilities of a set of single-dimensional MDPs for the order-holding problem.

2.3. Estimating Transition Probabilities of Markov Chains

Several methods exist in the literature to predict the transition probabilities of Markov chains. For Markov chains with finite discrete states, the MLE method (Anderson and Goodman 1957, Collins 1974) is often used to estimate the transition probabilities from samples. Specifically, the transition probability from state i to state j is estimated as $\hat{p}_{ij} = N_{ij}/\sum_{k=1}^{n} N_{ik}$, where N_{ik} represents the number of transitions from state i to state k in the training samples. However, this method fails to output the transition probabilities of newly observed states that are not in the training samples.

A targeting technology by Roussas (1969) and Yakowitz (1985) is developed to fix this problem. This technology shares the same idea as the kernel conditional density estimation method (Rosenblatt 1969, Bashtannyk and Hyndman 2001, Hall et al. 2004, Li and Racine 2007). A critical and often difficult step in the estimation is to choose unknown parameters called bandwidths. Li and Racine (2007) introduce two data-driven methods to choose bandwidths: a least-square cross-validation method and a maximum likelihood cross-validation method, which require solving non-convex programs and are computationally expensive. In contrast, our proposed estimation model is convex and much more tractable.

3. The Order-Holding Problem

We formulate the order-holding problem as an MDP in Section 3.1. Due to the curse of dimensionality, we do not solve this MDP directly. Instead, we first construct a set of single-dimensional MDPs in Section 3.2. In Section 3.3, we show that the optimal policy of the original MDP can be characterized by personalized thresholds that can be obtained by solving the set of single-dimensional MDPs, which are much more tractable.

3.1. MDP Formulation

We formulate an online retailer's order-holding problem for a single consumer as an MDP. We divide the time horizon into discrete periods. At the end of each period, the consumer decides whether to place an order. Define *dwell time* as the number of periods without new orders since the last order placed by the consumer. We assume that the consumer will leave the system immediately if she does not place any new order for T periods. That is, T is the maximum dwell time. We say two orders placed by the consumer are *consecutive* if the inter-arrival time between these two orders is no more than T periods.

A consumer's shopping journey begins with her first order, and ends when she leaves the system. The consumer immediately leaves the system if (i) she does not place any new order for T periods or (ii) she has placed L consecutive orders. During the consumer's shopping journey, the retailer can first hold several orders together as a *pending combined order* and then send the pending combined order to the 3PL at an appropriate time. Depending on the length of the consumer's shopping journey, the retailer may send multiple pending combined orders at the start of different periods during the shopping journey. Figure 2 shows an example of a shopping journey in which the retailer sends a combined order (containing the first and second orders) at the start of period 6 and another (containing only the third order) at the start of period 12. The consumer leaves the system at the end of period 14 and her shopping journey ends.

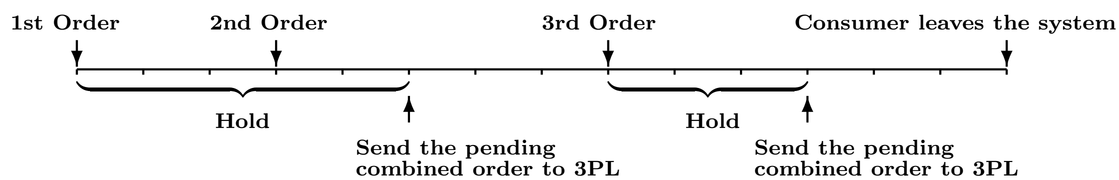


Figure 2 Example of a consumer's shopping journey

At the start of each period t, let l denote the number of *past consecutive orders* made by the consumer so far in her shopping journey. Let $\mathbf{x}_l$ denote the *order features* of the l past consecutive orders. These features include the consumer's demographic information and the product attributes. Specifically, the order features can be expressed as a column vector $\mathbf{x}_l = (\mathbf{x}_{l,0}, \mathbf{x}_{l,1}, \ldots, \mathbf{x}_{l,l})^{\intercal}$, where $\mathbf{x}_{l,k}$, $k = 1, \ldots, l$, represents the product attributes of the kth order among the past consecutive orders and $\mathbf{x}_{l,0}$ denotes other features including the consumer's demographic information. Let d_0 be the dimension of $\mathbf{x}_{l,0}$ and d_1 be the dimension of $\mathbf{x}_{l,k}$, $k = 1, \ldots, l$.

Define the state of the MDP model at the start of period t as $(\mathbf{x}_l, w, v)$. The second state variable w represents the dwell time since the latest order. The last state variable v equals 1 if there is a pending combined order or 0 otherwise. For example, in Figure 2 the state at the start of period 7 is $(\mathbf{x}_2, 3, 0)$. If the consumer places an order at the end of period t, then we set $w = 0$ and $v = 1$ at the start of period $t+1$. Since the consumer makes at most L consecutive orders, the state space is $\mathcal{S} = \{(\mathbf{x}_l, w, v) | \mathbf{x}_l \in X_l, 1 \le l \le L; w \in \{0, 1, \ldots, \infty\}; v \in \{0, 1\}\}$, where X_l represents a complete set of order features of the l past consecutive orders. At the start of each period t, the retailer first observes the system state and then decides whether to hold a pending combined order, if any, to period $t+1$. The combined order may include a single order or multiple consecutive orders made by the consumer earlier. The retailer's decision incurs a (holding or order arrangement) cost in period t. After that, the consumer may place a new order at the end of period t. Figure 3 shows the sequence of events in each period t.

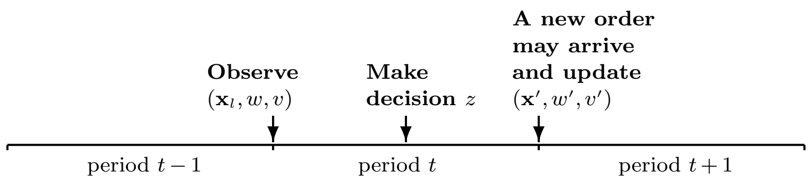


Figure 3 **The sequence of events in each period** t

If $v=1$ (there is a pending combined order) at the start of period t, the retailer has two possible actions: (i) Send the pending combined order to the 3PL immediately, denoted as $z=0$. (ii) Hold the pending combined order to the next period, denoted as $z=1$. The action $z=0$ incurs an *order arrangement cost* c, whereas the action $z=1$ incurs a *holding cost per period* h that represents the disutility in delivery time. Note that the order arrangement cost c represents a fixed cost to transmit a pending combined order to the 3PL. (In practice, the 3PL charges a fixed transaction fee for handling each combined order. This fixed fee does not include the actual order-picking cost in the warehouse, which is a variable cost that does not affect the order-holding decisions.) We assume that c and h are independent of the number of items in the combined order. As shown in Lemma 2 in Appendix A, it is suboptimal to hold some orders, while sending other received orders to the 3PL. That is, the pending combined order should be sent as a whole if the retailer decides to send it. If $v=0$ at the start of period t, the retailer does nothing, denoted as $z=2$, and no cost is incurred in the period.

For convenience, we use the term "state" to refer to $(\mathbf{x}_l, w, v)$ or each of the variables $\mathbf{x}_l$, w, v, and their combinations. Given the state $(\mathbf{x}_l, w)$ at the start of period t, we now describe its evolution. If no new order arrives at the end of period t, the order features $\mathbf{x}_l$ do not change and the next state to visit is $(\mathbf{x}_l, w+1)$. Otherwise, if a new order with product attributes $\hat{\boldsymbol{x}}$ arrives at the end of period t, the next state to visit is $(\mathbf{x}_{l+1}, 0)$, where $\mathbf{x}_{l+1} = (\mathbf{x}_l{}^{\intercal}, \hat{\boldsymbol{x}})^{\intercal}$. Thus, the set of next states to visit can be determined as $\sigma((\mathbf{x}_l, w)) = \{(\mathbf{x}_l, w+1)\} \cup \{(\mathbf{x}_{l+1}, 0) | \mathbf{x}_{l+1} \in X_{l+1}\}$. The state $(\mathbf{x}_l, T)$ means that no new order arrives for T periods (the maximum dwell time) after the arrival of the lth order. In this case, the consumer will leave the system immediately and the next state to visit is $(\mathbf{x}_l, T+1)$. That is, $\sigma((\mathbf{x}_l, T)) = \{(\mathbf{x}_l, T+1)\}$. Define $p((\mathbf{x}_l, w), (\mathbf{x}', w'))$ as the transition probability from the state $(\mathbf{x}_l, w)$ to the next state $(\mathbf{x}', w') \in \sigma((\mathbf{x}_l, w))$. We have $p((\mathbf{x}_l, w), (\mathbf{x}_l, w+1)) = 1$ for

$w \geq T$. Similarly, we have $p((\mathbf{x}_L, w), (\mathbf{x}_L, w+1)) = 1$ for $w \geq 0$ because the consumer makes at most L consecutive orders, after which she will leave the system immediately.

Recall the sequence of events in Figure 3. Given the retailer's action z in period t, let $P((\mathbf{x}_l, w, v), (\mathbf{x}', w', v'); z)$ denote the transition probability from state $(\mathbf{x}_l, w, v)$ to state $(\mathbf{x}', w', v')$. Define $\mathcal{I}(\cdot)$ as an indicator function such that $\mathcal{I}(\cdot)$ equals 1 if $\cdot$ is true, and 0 otherwise. Thus, $\mathcal{I}((\mathbf{x}', w') \neq (\mathbf{x}_l, w+1))$ indicates whether there is a new order arriving at the end of period t. If $v = 1$ at the start of period t, for any $(\mathbf{x}', w') \in \sigma((\mathbf{x}_l, w))$, we have the state dynamics:

$$\begin{aligned} P((\mathbf{x}_l, w, v=1), (\mathbf{x}', w', v' = \mathcal{I}((\mathbf{x}', w') \neq (\mathbf{x}_l, w+1))); z=0) &= p((\mathbf{x}_l, w), (\mathbf{x}', w')), \\ P\big((\mathbf{x}_l, w, v=1), (\mathbf{x}', w', v'=1); z=1\big) &= p((\mathbf{x}_l, w), (\mathbf{x}', w')). \end{aligned} \tag{1}$$

The first equation of (1) means that if the retailer sends the pending combined order to the 3PL $(z=0)$ in period t, then we set v' depending on whether there is a pending combined order emerging. That is, $v' = 1$ if a new order arrives at the end of period t or $v' = 0$ otherwise. The second equation of (1) means that if the retailer continues to hold the pending combined order $(z=1)$ in period t, then we set $v' = 1$. In contrast, if $v = 0$ at the start of period t, then the retailer does nothing $(z=2)$ in period t, and we have the state dynamics:

$$P((\mathbf{x}_l, w, v=0), (\mathbf{x}', w', v' = \mathcal{I}((\mathbf{x}', w') \neq (\mathbf{x}_l, w+1))); z=2) = p((\mathbf{x}_l, w), (\mathbf{x}', w')). \tag{2}$$

That is, we set v' depending on whether there is a pending combined order emerging. Note that in (1) and (2), the retailer's action z can only affect v' but not $(\mathbf{x}', w')$.

Given state $(\mathbf{x}_l, w, v)$ at the start of period t, define $V(\mathbf{x}_l, w, v)$ as the optimal total expected cost from period t until the consumer leaves the system. The Bellman equations are

$$\begin{aligned} V(\mathbf{x}_l, w, 1) = \min_{z \in \{0,1\}} z &\left[h + \sum_{(\mathbf{x}', w') \in \sigma((\mathbf{x}_l, w))} p((\mathbf{x}_l, w), (\mathbf{x}', w')) \cdot V(\mathbf{x}', w', 1) \right] \\ + (1-z) &\left[c + \sum_{(\mathbf{x}', w') \in \sigma((\mathbf{x}_l, w))} p((\mathbf{x}_l, w), (\mathbf{x}', w')) \cdot V(\mathbf{x}', w', v' = \mathcal{I}((\mathbf{x}', w') \neq (\mathbf{x}_l, w+1))) \right], \end{aligned} \tag{3}$$

$$V(\mathbf{x}_l, w, 0) = \sum_{(\mathbf{x}', w') \in \sigma((\mathbf{x}_l, w))} p((\mathbf{x}_l, w), (\mathbf{x}', w')) \cdot V(\mathbf{x}', w', v' = \mathcal{I}((\mathbf{x}', w') \neq (\mathbf{x}_l, w+1))). \tag{4}$$

$V(\mathbf{x}_l, w, v)$ includes the costs incurred in period t and the expected cost to go. Unfortunately, the above MDP has an infinite number of states because the variable w can be any nonnegative integer. Note that there are two types of states that imply the consumer leaves the system immediately. The first is $(\mathbf{x}_l, T, v)$ meaning that no new order arrives during the previous T periods. The second is $(\mathbf{x}_L, 0, 1)$ meaning that the consumer has placed the maximum number of consecutive orders. We find the retailer's optimal actions for these states in Lemma 1.

Lemma 1 (Terminal Conditions). *For state* $(\mathbf{x}_l, T, v)$ *or* $(\mathbf{x}_L, 0, 1)$ *at the start of period* t*, the consumer will leave the system and it is optimal to send the pending combined order, if any, to the 3PL immediately. In addition, we have* $V(\mathbf{x}_l, T, 0) = 0$*,* $V(\mathbf{x}_l, T, 1) = c$*, for* $l \le L-1$*; and* $V(\mathbf{x}_L, 0, 1) = c$*.*

Lemma 1 reduces the MDP (3-4) to a finite number of states as follows. For notational convenience, let $\langle n\rangle = \{0, 1, \ldots, n\}$ denote a set of integers for any n. For $l \le L-1$ and $w \in \langle T-1\rangle$,

$$V(\mathbf{x}_l, w, 1) = \min_{z\in\{0,1\}} z\left[h + \sum_{(\mathbf{x}',w')\in\sigma((\mathbf{x}_l,w))} p((\mathbf{x}_l, w), (\mathbf{x}', w'))\cdot V(\mathbf{x}', w', 1)\right]$$
$$+ (1-z)\left[c + \sum_{(\mathbf{x}',w')\in\sigma((\mathbf{x}_l,w))} p((\mathbf{x}_l, w), (\mathbf{x}', w'))\cdot V\left(\mathbf{x}', w', v' = \mathcal{I}((\mathbf{x}', w') \ne (\mathbf{x}_l, w+1))\right)\right], \tag{5}$$

$$V(\mathbf{x}_l, w, 0) = \sum_{(\mathbf{x}',w')\in\sigma((\mathbf{x}_l,w))} p((\mathbf{x}_l, w), (\mathbf{x}', w'))\cdot V\left(\mathbf{x}', w', v' = \mathcal{I}((\mathbf{x}', w') \ne (\mathbf{x}_l, w+1))\right), \tag{6}$$

$$V(\mathbf{x}_l, T, 0) = 0, \tag{7}$$

$$V(\mathbf{x}_l, T, 1) = c, \tag{8}$$

$$V(\mathbf{x}_L, 0, 1) = c. \tag{9}$$

The number of states in (5-9) is linked to the retailer's number of stock-keeping units (SKUs), which can be huge in practice, making the MDP still intractable. Next, we will construct a set of single-dimensional MDPs and show that we can obtain the optimal policy of the original MDP by solving the single-dimensional MDPs. In this way, we can overcome the curse of dimensionality.

3.2. Overcoming The Curse of Dimensionality

We first construct a set of single-dimensional MDPs and then show that their optimal decisions coincide with that of the original MDP (5-9). We group the states $(\mathbf{x}_l, w)$ into clusters, with each cluster identified by $\mathbf{x}_l$. Specifically, we define a cluster $\mathfrak{C}_{\mathbf{x}_l} = \{(\mathbf{x}_l, w) | 0 \le w \le T\}$ for $\mathbf{x}_l \in X_l$ and $l \le L-1$. Given $\mathbf{x}_l$ at the start of period t, define a *single-dimensional MDP* as follows:

$$\Delta_{\mathbf{x}_l}(w) = \min_{z\in\{0,1\}} z\left[h + p((\mathbf{x}_l, w), (\mathbf{x}_l, w+1))\cdot \Delta_{\mathbf{x}_l}(w+1)\right] + (1-z)c, \quad 0 \le w \le T-1; \tag{10}$$
$$\Delta_{\mathbf{x}_l}(T) = c;$$

where $\Delta_{\mathbf{x}_l}(w)$ denotes the optimal total expected cost from state w to state T. We associate $\Delta_{\mathbf{x}_l}(\cdot)$ with $\mathbf{x}_l$ to emphasize the dependence of the single-dimensional MDP (10) on the order features. When the consumer places a new order with product attributes $\hat{\boldsymbol{x}}$, the system transits from a single-dimensional MDP associated with $\mathbf{x}_l$ to another single-dimensional MDP associated with

$\mathbf{x}_{l+1} = (\mathbf{x}_l{}^\intercal, \hat{\boldsymbol{x}})^\intercal$. We will model this transition using a consumer's sequential decision process in Section 4.

We next show that the single-dimensional MDP defined for the cluster $\mathfrak{C}_{\mathbf{x}_l}$ shares the same optimal decisions for the states related to $\mathbf{x}_l$ in the original MDP (5-9).

Proposition 1. *For each cluster $\mathfrak{C}_{\mathbf{x}_l}$, given state $(\mathbf{x}_l, w, 1)$ with $(\mathbf{x}_l, w) \in \mathfrak{C}_{\mathbf{x}_l}$, there exists an optimal decision $z^*(\mathbf{x}_l, w, 1)$ for state $(\mathbf{x}_l, w, 1)$ in the original MDP (5-9) that is equivalent to an optimal decision $z^*_{\mathbf{x}_l}(w)$ for state w in the single-dimensional MDP (10) associated with $\mathbf{x}_l$.*

Proposition 1 implies that instead of solving the original MDP (5-9), we can solve a set of single-dimensional MDPs (10) with the same solution structure. The state space of each single-dimensional MDP only depends on the maximum dwell time T rather than the number of SKUs, which is often huge for online retailing. *This significantly reduces the computation burden.* This approach not only overcomes the curse of dimensionality, but also admits a *personalized* order-holding policy because each single-dimensional MDP is specified by order features $\mathbf{x}_l$. We first show how to characterize the optimal policy for MDP (5-9) based on the single-dimensional MDPs (10).

3.3. Structure of The Optimal Policy

Theorem 1 characterizes the optimal order-holding policy for MDP (5-9).

Theorem 1 (Optimal Order-Holding Policy).
1. For past consecutive orders featured by $\mathbf{x}_l$, there exists a threshold $S(\mathbf{x}_l)$ such that the optimal order-holding policy for MDP (5-9) is as follows: If no new order arrives for $S(\mathbf{x}_l)$ periods since the latest order, send the pending combined order to the 3PL; otherwise add the newly received order with product attributes $\hat{\boldsymbol{x}}$ to the pending combined order and update the threshold as $S(\mathbf{x}_{l+1})$, where $\mathbf{x}_{l+1} = (\mathbf{x}_l{}^\intercal, \hat{\boldsymbol{x}})^\intercal$.
2. The optimal threshold $S(\mathbf{x}_l)$ can be obtained from the single-dimensional MDP (10) as

$$S(\mathbf{x}_l) = \min\left\{\left\{w \in \langle T-1\rangle | h + p((\mathbf{x}_l, w), (\mathbf{x}_l, w+1)) \cdot \Delta_{\mathbf{x}_l}(w+1) \geq c\right\} \cup \{T\}\right\}. \tag{11}$$

It is straightforward to implement the threshold-type policy in Theorem 1 as follows. For each newly received order, the online retailer can determine the optimal threshold by the order features from (11). The retailer should hold the pending combined order if $w < S(\mathbf{x}_l)$, but send it to the 3PL if $w \geq S(\mathbf{x}_l)$. When a new order with product attributes $\hat{\boldsymbol{x}}$ arrives, an *order evolution* occurs: We update the features of the past consecutive orders as $\mathbf{x}_{l+1} = (\mathbf{x}_l{}^\intercal, \hat{\boldsymbol{x}})^\intercal$ and the threshold as $S(\mathbf{x}_{l+1})$, which can be calculated from (11). (Note that under the framework of the single-dimensional MDPs (10), the order evolution corresponds to a transition from a single-dimensional MDP associated with $\mathbf{x}_l$ to another single-dimensional MDP associated with $\mathbf{x}_{l+1}$.) This procedure repeats until no new order arrives for T periods (the maximum dwell time) or the consumer places her Lth order, upon which she will leave the system immediately.

4. Consumer's Sequential Decision Model

To model the transitions among the single-dimensional MDPs, we introduce a choice model to characterize a consumer's sequential decision process after placing an order. We will use the choice model to predict the consumer's probability of placing consecutive orders and the transition probabilities in each single-dimensional MDP (10) as well as among the different single-dimensional MDPs. This parametric model has several advantages. First, it is easy to interpret as it characterizes how a consumer chooses among three alternatives: search and see, leave the system, or place another order. Second, it mitigates the overfitting issue as it has much fewer parameters to estimate compared to non-parametric models. Finally, it is computationally efficient. Based on this choice model, we express in closed form the transition probabilities in order features, allowing us to analytically characterize how the order features affect the optimal threshold.

4.1. Consumer's Sequential Decision Process

To determine the optimal threshold in Theorem 1, we need to predict the transition probabilities in each single-dimensional MDP (10). We adopt a sequential choice model for the consumer's decision process after she places an order. We then relate the choice probabilities of this model to the transition probabilities in each of and among the single-dimensional MDPs. Based on the 2020 MSOM Data Driven Research Challenge data, only 8.85% of the orders request multiple items. Thus, we assume at most one item purchased per order and the order features $\mathbf{x}$ involve *only single-item orders*. We can split a multi-item order into consecutive orders to fit this assumption.

After placing an order, the consumer has the following three options: (i) dwell on the platform to search further, (ii) leave the system immediately, or (iii) place another order. We define a directed network $N_{\mathbf{x}_l} = (\mathcal{V}_{\mathbf{x}_l} \cup \{\triangle, M\}, \mathcal{A}_{\mathbf{x}_l})$ to model this sequential decision process, where $\mathcal{V}_{\mathbf{x}_l}$ represents a set of nodes, $\triangle$ represents the option of leaving the system immediately, M represents the option of placing another order, and $\mathcal{A}_{\mathbf{x}_l}$ represents a set of arcs. Each network is associated with the order features $\mathbf{x}_l$ of the past consecutive orders placed by the consumer in her shopping journey. Figure 4 illustrates the consumer's sequential decision process. Each node in $\mathcal{V}_{\mathbf{x}_l}$ corresponds to a state $(\mathbf{x}_l, w)$ defined in Section 3.1. From each node $(\mathbf{x}_l, w)$, there are arcs in $\mathcal{A}_{\mathbf{x}_l}$ linking the node to three possible successors corresponding to the above three options.

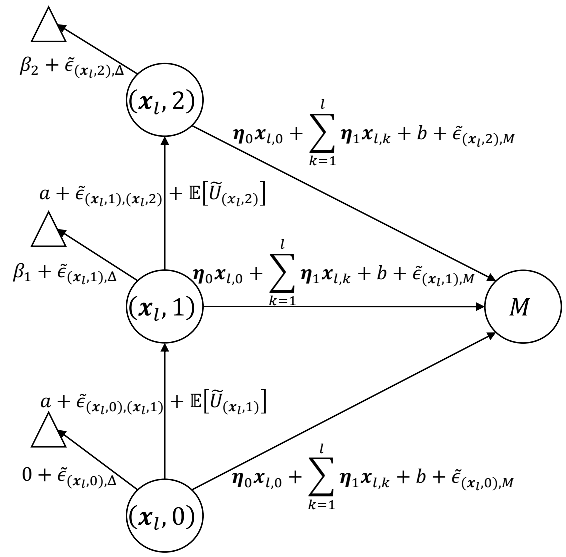


Figure 4 **A choice network with order features $\mathbf{x}_l$ and maximum dwell time $T = 3$**

To determine the choice probabilities, we model the utility function of each option (dwell on the platform, leave the system immediately, or place another order). Following Train (2009) on discrete choice models, we assume an additive structure in the utility function. Specifically, from the current node $(\mathbf{x}_l, w)$, the utility of choosing node k consists of a deterministic term $u_{(\mathbf{x}_l,w),k}$ and a random error term $\tilde{\epsilon}_{(\mathbf{x}_l,w),k}$. We model the utility function in a Markovian manner as follows:

$$\mathbb{E}\left[\tilde{U}_{(\mathbf{x}_l,w)}\right] = \mathbb{E}\left[\max_{k:((\mathbf{x}_l,w),k)\in\mathcal{A}_{\mathbf{x}_l}} \left\{u_{(\mathbf{x}_l,w),k} + \tilde{\epsilon}_{(\mathbf{x}_l,w),k} + \mathbb{E}\left[\tilde{U}_k\right]\right\}\right], \text{ for } (\mathbf{x}_l, w) \in \mathcal{V}_{\mathbf{x}_l}; \tag{12}$$
$$\mathbb{E}\left[\tilde{U}_k\right] = 0, \text{ for } k \in \{\triangle, M\};$$

where $\mathbb{E}\left[\tilde{U}_k\right]$ is the expected maximum utility of node k. The consumer's utility of choosing an option in (12) is the sum of its immediate utility and the expected utility of the subsequent options.

We model the deterministic utility term as a linear function of features. Specifically, we assume $u_{(\mathbf{x}_l,w),(\mathbf{x}_l,w+1)} = a$ (≤ 0), for $w \in \langle T-2 \rangle$, to represent the consumer's searching cost if she decides to dwell and search for another period. We model the consumer's deterministic utility when she chooses to leave the system as $u_{(\mathbf{x}_l,w),\triangle} = \beta_w$, where $\beta_w \in \mathbb{R}$ for $w \in \langle T-1 \rangle$. Without loss of generality, we normalize β_0 to zero. The third option for each node is to place another order (by choosing node M), which triggers an order evolution. When the consumer places another order with product attributes $\hat{\boldsymbol{x}}$, the system transits from a single-dimensional MDP associated with $\mathbf{x}_l$ to another single-dimensional MDP associated with $\mathbf{x}_{l+1} = (\mathbf{x}_l{}^{\mathsf{T}}, \hat{\boldsymbol{x}})^{\mathsf{T}}$. Note that $\mathbf{x}_l$ includes the consumer's demographic information and her past consecutive orders' product attributes. Recall that d_0 represents the dimension of $\mathbf{x}_{l,0}$ and d_1 represents the dimension of $\mathbf{x}_{l,k}$, $k=1,\dots,l$. To construct this option's utility, define $u_{(\mathbf{x}_l,w),M} = \boldsymbol{\eta}_0\mathbf{x}_{l,0} + \sum_{k=1}^{l} \boldsymbol{\eta}_1\mathbf{x}_{l,k} + b$, where $\boldsymbol{\eta}_0 \in \mathbb{R}^{d_0}$, $\boldsymbol{\eta}_1 \in \mathbb{R}^{d_1}$, and $b \in \mathbb{R}$, as the deterministic utility of placing another order. In this function, the marginal effect of each attribute on the utility of purchasing another item is the same across all the past consecutive orders. The incentive to place another order depends on the overall product bundle of the past consecutive orders, but not the sequence of these orders. For notational convenience, we define $\boldsymbol{\eta} = (\boldsymbol{\eta}_0, \boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_1)$ and write $u_{(\mathbf{x}_l,w),M} = \boldsymbol{\eta}\mathbf{x}_l + b$. This utility model captures consumer and purchased-product heterogeneity through $\mathbf{x}_l$, which vary across consumers and their sets of past consecutive orders.

For convenience, we drop the subscript $\mathbf{x}_l$ for the rest of this section, and simply denote the set of nodes, the set of arcs, and the network as $\mathbf{V} = \{0,\dots,w-1\} \cup \{\triangle, M\}$, $\mathbf{A} = \{(w,w+1)|w \in \langle T-2\rangle\} \cup \{(w,\triangle)|w \in \langle T-1\rangle\} \cup \{(w,M)|w \in \langle T-1\rangle\}$, and $N = (\mathbf{V},\mathbf{A})$ respectively. Denote a collection of coefficients $a, b, \boldsymbol{\eta}$, and $\boldsymbol{\beta} = (\beta_0,\dots,\beta_{T-1})$ as $\Theta = (a,b,\boldsymbol{\eta},\boldsymbol{\beta})$. We assume the random error terms $\tilde{\epsilon}_{(\mathbf{x}_l,w),k}$ in (12) follow i.i.d. Gumbel distributions. We characterize the choice probability using a *Markovian logit model* (MLM) (Bell 1995). Under MLM, the *conditional* choice probability $\bar{p}_{w,k}(\mathbf{x}_l;\Theta,N)$ of option k from node w can be written as follows:

$$
\begin{aligned}
\bar{p}_{w,w+1}(\mathbf{x}_l;\Theta,N) &= \frac{\exp(a + g_{w+1}(\mathbf{x}_l;\Theta,N))}{\exp(g_w(\mathbf{x}_l;\Theta,N))}, \text{ for } w \in \langle T-2\rangle;\\
\bar{p}_{w,\triangle}(\mathbf{x}_l;\Theta,N) &= \frac{\exp(\beta_w)}{\exp(g_w(\mathbf{x}_l;\Theta,N))}, \text{ for } w \in \langle T-1\rangle;\\
\bar{p}_{w,M}(\mathbf{x}_l;\Theta,N) &= \frac{\exp(\boldsymbol{\eta}\mathbf{x}_l + b)}{\exp(g_w(\mathbf{x}_l;\Theta,N))}, \text{ for } w \in \langle T-1\rangle; \text{ where}\\
g_w(\mathbf{x}_l;\Theta,N) &= \ln\left[\exp(a + g_{w+1}(\mathbf{x}_l;\Theta,N)) + \exp(\beta_w) + \exp(\boldsymbol{\eta}\mathbf{x}_l + b)\right], \text{ for } w \in \langle T-2\rangle;\\
g_{T-1}(\mathbf{x}_l;\Theta,N) &= \ln\left[\exp(\beta_{T-1}) + \exp(\boldsymbol{\eta}\mathbf{x}_l + b)\right].
\end{aligned}
\tag{13}
$$

4.2. Determining the Consumer's Probability of Placing Consecutive Orders and the Transition Probabilities in the Single-Dimensional MDP

At the end of each period t, the consumer has three options: dwelling, leaving, and placing another order on the choice network $N = (\mathbf{V},\mathbf{A})$. After she places her latest order, a consecutive order occurs if it is placed within T periods. Given the features $\mathbf{x}_l$ of her past consecutive orders, the consumer's *probability of placing consecutive orders* is equal to the sum of the probabilities for her to place another order within T periods: $\sum_{w=0}^{T-1} \bar{q}_{w,M}(\mathbf{x}_l;\Theta,N)$, where $\bar{q}_{w,M}(\mathbf{x}_l;\Theta,N)$ represents the *unconditional* probability of arc $(w,M)\in\mathbf{A}$, which depends on the choice network N, coefficients Θ, and order features $\mathbf{x}_l$. The unconditional choice probability of any arc $(w,k)\in\mathbf{A}$ is

$$\bar{q}_{w,k}(\mathbf{x}_l;\Theta,N)=\prod_{(n_1,n_2)\in route(0,w,k)}\bar{p}_{n_1,n_2}(\mathbf{x}_l;\Theta,N), \tag{14}$$

where $route(0,w,k)$ represents a unique route from node 0, through node w, to node k in the network $N=(\mathbf{V},\mathbf{A})$, and (n_1,n_2) represents an arc on the route. Note that we find the unconditional probability $\bar{q}_{w,k}$ using a series of conditional probabilities $\bar{p}_{n_1,n_2}$ in (14).

Proposition 2 (Probability of Placing Consecutive Orders). *Under MLM (13) with network N, coefficients Θ, and order features $\mathbf{x}_l$, the probability of placing consecutive orders is $\frac{1-e^{aT}}{1-e^{a}}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)$.*

Proposition 2 provides an analytical expression that helps the online retailer predict the probability for the consumer to place a consecutive order after her latest order. Note that if the consumer places a consecutive order, the system transits from the single-dimensional MDP associated with $\mathbf{x}_l$ to another single-dimensional MDP. Thus, the probability of placing consecutive orders also characterizes the transition probability among the single-dimensional MDPs (10).

After receiving the latest order, the retailer will not observe any new orders in w periods if the consumer leaves the system within or dwells for at least w periods. The probability of having no new orders within w periods is

$$\mathbb{P}(w;\mathbf{x}_l,\Theta,N)=\begin{cases}1, \text{ for } w=0,\\ \bar{q}_{w-1,w}(\mathbf{x}_l;\Theta,N)+\sum_{\tau=0}^{w-1}\bar{q}_{\tau,\triangle}(\mathbf{x}_l;\Theta,N), \text{ for } 1\le w\le T-1,\\ \sum_{\tau=0}^{T-1}\bar{q}_{\tau,\triangle}(\mathbf{x}_l;\Theta,N), \text{ for } w=T.\end{cases} \tag{15}$$

We find the transition probability $p((\mathbf{x}_l,w),(\mathbf{x}_l,w+1))$ through the conditional probability formula:

$$p((\mathbf{x}_l,w),(\mathbf{x}_l,w+1))=\mathbb{P}(w+1;\mathbf{x}_l,\Theta,N)/\mathbb{P}(w;\mathbf{x}_l,\Theta,N), \text{ for } w\in\langle T-1\rangle. \tag{16}$$

Together with (13-15), (16) establishes a closed-form relation between the order features $\mathbf{x}_l$ and the transition probabilities in each single-dimensional MDP (10).

4.3. Determining The Optimal Personalized Threshold

Theorem 1 shows that the optimal holding threshold can be obtained from the single-dimensional MDP (10), which is determined by the transition probabilities in (16). Based on this closed-form relation between the transition probabilities and the order features, we further express the optimal threshold in order features analytically. Given $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)$, define a nonlinear function

$$f_{nl}(\bar{q}_{0,M})=\frac{1}{a}\left\{\ln(h)-\ln\left(c-\frac{h}{1-e^{a}}\right)\right\}-\frac{1}{a}\ln(\bar{q}_{0,M})+\frac{1}{a}\ln\left(1-\frac{1}{1-e^{a}}\bar{q}_{0,M}\right). \tag{17}$$

Define $A=T$ if $W_t-c<0$ for all $t\in\langle T-1\rangle$, and $A=0$ otherwise; where

$$W_t = h+\left(1-\frac{e^{at}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)}{1-\frac{1-e^{at}}{1-e^{a}}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)}\right)\cdot W_{t+1},\ \text{for } t\in\langle T-1\rangle,$$
$$W_T = c.$$

The following theorem determines the personalized threshold based on the closed-form relation (16) between the transition probabilities of the single-dimensional MDP (10) and the order features.

Theorem 2 (The Optimal Personalized Threshold). *Under MLM (13) with network N, coefficients Θ, and order features $\mathbf{x}_l$, the optimal personalized threshold is determined as follows: If $h<c\cdot(1-e^a)$,*

$$S(\mathbf{x}_l;\Theta,N)=\begin{cases}(\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0, & \text{if } \bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)<1-e^a; & \text{(18a)}\\ T, & \text{otherwise.} & \text{(18b)}\end{cases}$$

Otherwise,

$$S(\mathbf{x}_l;\Theta,N)=\begin{cases}0, & \text{if } \bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\le 1-e^a; & \text{(19a)}\\ A, & \text{if } 1-e^a<\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\le h/c; & \text{(19b)}\\ T, & \text{if } \bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)>h/c. & \text{(19c)}\end{cases}$$

The closed-form expression of the personalized threshold in Theorem 2 characterizes its dependence on the order features. Specifically, the order features $\mathbf{x}_l$ leading to a higher probability of placing consecutive orders ($\frac{1-e^{aT}}{1-e^a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)$) correspond to a higher threshold. Based on this observation, we further explore a monotonic relationship between the order features and the threshold. For notational convenience, let $[n]=\{1,\ldots,n\}$ denote a set of integers for any n.

Corollary 1 (Monotonicity). *Under MLM (13), the following properties hold:*

1. *The optimal threshold is increasing in the probability of placing consecutive orders.*
2. *For a numerical feature x and its corresponding coefficient η, both the probability of placing consecutive orders and the optimal threshold are increasing (decreasing) in x if $\eta>0$ ($\eta<0$).*

The proof of Corollary 1 is straightforward based on the analytical expressions in (14) and Theorem 2 and is omitted. Part 1 of Corollary 1 shows that the optimal threshold is positively correlated with the consumer's probability of placing consecutive orders. Part 2 characterizes the effect of the order features $\mathbf{x}_l$ on the threshold $S(\mathbf{x}_l;\Theta,N)$. For a *numerical* feature x, if the coefficient η is positive, the probability $\frac{1-e^{aT}}{1-e^a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)$ of placing a consecutive order and the optimal threshold increase with x. Corollary 1 provides managers a guideline to select products with desired attributes (when recommending products) in order to control the probability of placing consecutive orders. We will discuss the effect of *categorical* features on the threshold in Section 7.2.

It is hard to analyze the marginal effect of the features $\mathbf{x}_l$ on the threshold $S(\mathbf{x}_l;\Theta,N)$ in Theorem 2 as it involves a nonlinear function $f_{nl}(\bar{q}_{0,M})$. To better visualize the relation between the order features and the threshold, we provide a piecewise-linear approximation of the threshold together with its performance guarantee below. Define a linear function $f_l(\mathbf{x}_l;\Theta)$ and a scalar ϵ_U as

$$f_l(\mathbf{x}_l;\Theta) = -\frac{1}{a}(\boldsymbol{\eta}\mathbf{x}_l + b) + C_0, \tag{20}$$

$$\epsilon_U = \frac{\frac{1-e^{aT}}{1-e^a}\exp(-a(T-1-C_0))}{\sum_{\tau\in\langle T-1\rangle}\exp(\beta_\tau+\tau a) + \frac{1-e^{aT}}{1-e^a}\exp(-a(T-1-C_0))}, \tag{21}$$

where $C_0 = \frac{1}{a}\left[\ln(h) - \ln\left(c - \frac{h}{1-e^a}\right) + \ln\left(\sum_{\tau\in\langle T-1\rangle}\exp(\beta_\tau+\tau a)\right)\right]$.

Theorem 3 (The Uniform Bound of Piecewise-Linear Approximation). *Under MLM (13) with network N, coefficients Θ, and order features $\mathbf{x}_l$, if $h < c(1-e^a)/(1+e^a)$, then for any $\mathbf{x}_l \in X_l$, $S(\mathbf{x}_l;\Theta,N)$ can be uniformly bounded by a piecewise-linear function as follows:*

$$0 \le S(\mathbf{x}_l;\Theta,N) - (\lceil f_l(\mathbf{x}_l;\Theta)\rceil \wedge T)\vee 0 \le \left\lfloor 1 + \frac{1}{a}\ln\left(1+\frac{\epsilon_U}{1-\epsilon_U}\right) + \frac{1}{a}\ln\left(1-\frac{\epsilon_U}{1-e^{aT}}\right)\right\rfloor. \tag{22}$$

Theorem 3 identifies a piecewise-linear approximation $f_{pl}(\mathbf{x}_l;\Theta) := (\lceil f_l(\mathbf{x}_l;\Theta)\rceil \wedge T)\vee 0$ to the personalized threshold $S(\mathbf{x}_l;\Theta,N)$. The approximation gap is upper bounded by $B(\epsilon_U,a) := \left\lfloor 1 + \frac{1}{a}\ln\left(1+\frac{\epsilon_U}{1-\epsilon_U}\right) + \frac{1}{a}\ln\left(1-\frac{\epsilon_U}{1-e^{aT}}\right)\right\rfloor$, which is shown to be in $O(1)$ by the numerical studies in Section 7.2. The piecewise-linear function with an approximation guarantee in Theorem 3 allows us to see the marginal effect of each feature on the threshold and hence identify the important features. We will discuss this in detail in Section 7.2. Note that the cost condition $h < c(1-e^a)/(1+e^a)$ for Theorem 3 is stricter than that for Theorem 2. In Proposition 4 in Appendix A, we relax this cost condition by adding a condition on the probability of placing consecutive orders, and derive a similar piecewise-linear approximation.

5. Estimating coefficients

We apply the maximum likelihood estimation (MLE) method to obtain the coefficients $\Theta = (a, b, \boldsymbol{\eta}, \boldsymbol{\beta})$ of MLM (13) using the data of 2020 MSOM Data Driven Research Challenge. The data set contains transaction records of the orders of each consumer and her click information.

Each *transaction path* begins from the first order placed by a consumer and terminates with an order after which she leaves the system. After an order is placed, we can observe from the data whether the consumer places another order within T periods (the maximum dwell time). If so, the dwell time is obtained as the time difference between the two consecutive orders, and we add the new order to the transaction path with l incremented by one. If not, we use the time point of the last click within the T periods to approximate the consumer's departure time from the system.

In this case, the dwell time is obtained as the time difference between the latest order and the consumer's departure. (For consumers without any click information in the data, we assume they leave the system at the end of the first period after they place an order. That is, the dwell time equals 1.) Let L_k denote the number of orders in transaction path k. Let $path^k = \{(\mathbf{x}_l^k, DT_l^k)\}_{l\in[L_k]}$ denote transaction path k, where $\mathbf{x}_l^k$ represents the features of the first l past consecutive orders in path k and DT_l^k is the dwell time since the lth order in path k.

We are now ready to construct the likelihood function of MLE. The likelihood of observing the lth entry $(\mathbf{x}_l^k, DT_l^k)$ of $path^k$ can be calculated as $\Pr\left(\mathbf{x}_l^k, DT_l^k;\Theta\right) = \prod_{w=0}^{DT_l^k-2} \bar{p}_{w,w+1}\left(\mathbf{x}_l^k;\Theta,N\right) \cdot \bar{p}_{DT_l^k-1,\Delta}\left(\mathbf{x}_l^k;\Theta,N\right)^{y_l^k} \cdot \bar{p}_{DT_l^k-1,M}\left(\mathbf{x}_l^k;\Theta,N\right)^{1-y_l^k}$, where y_l^k is an indicator variable that equals 1 if the consumer leaves the system after the lth order or 0 otherwise. For convenience, define $\prod_{w=0}^{-1} \bar{p}_{w,w+1}\left(\mathbf{x}_l^k;\Theta,N\right) = 1$. Let K denote the number of paths in the data set. By applying (13), we can formulate the MLE model (Train 2009) as follows:

$$\hat{\Theta} = \arg \max_{\Theta=(a,b,\boldsymbol{\eta},\boldsymbol{\beta})} \sum_{k=1}^{K}\sum_{l=1}^{L_k} \ln\left(\Pr\left(\mathbf{x}_l^k, DT_l^k;\Theta\right)\right) \qquad (23)$$
$$s.t.\ a \le 0.$$

Proposition 3 (Convexity of The MLE Model). *The MLE model (23) is a convex program with respect to* $\Theta = (a, b, \boldsymbol{\eta}, \boldsymbol{\beta})$.

Proposition 3 implies that the MLE model (23) can be efficiently solved. Lemma 7 in Appendix A shows that the MLE method applied to MLM above is equivalent to the logistic regression model (Ng 2004). This implies that the estimation model (23) provides a probabilistic classifier and shares the advantages of the logistic regression model. For example, it predicts a well-calibrated probability distribution over a set of classes (Niculescu-Mizil and Caruana 2005), rather than outputs only the most-likely class that the observation should belong to.

6. Data Processing

We present a procedure to extract each transaction path defined in Section 5 from the data of 2020 MSOM Data Driven Research Challenge (Shen et al. 2019), which contains transaction-level data from JD.com over one month. Table 9 in Appendix G shows the feature information of 31,868 SKUs such as *type* (indicating whether the SKU is sold by JD), *brand ID*, *attribute1*, and *attribute2*. Table 10 provides the demographic information of 457,298 consumers such as *gender*, *age*, *marital status*, *education*, *user level* (indicating their historical purchase value) and *plus* (indicating whether the consumer is a JD plus member). Table 11 shows detailed transaction data of 486,928 orders such as *order time*, *original unit price*, *final price*, and other discount details. Table 12 lists 17,906,095 click actions from 2,557,836 consumers showing who clicks which item at what time. Note that 31.13% of the orders are placed by users (consumers) with missing demographic information. The proportion is even up to 99.66% for the orders made by enterprise users. Due to this reason, we keep all the enterprise users regardless of their demographic information. For other user levels, we drop their orders if their demographic information has missing values, which make up 30.06% of all the orders.

To apply the threshold-type policy, we need to estimate the coefficients $\Theta = (a, b, \boldsymbol{\eta}, \boldsymbol{\beta})$ as described in Section 5, which requires extracting the transaction paths $\{path^k\}_{k\in[K]}$ from the data. Recall that the lth entry of $path^k = \{(\mathbf{x}_l^k, DT_l^k)\}_{l\in[L_k]}$ is specified by the order features and the dwell time since the lth order. We set each period as one minute, and so the maximum dwell time is T minutes. We use the order and click information from Tables 11-12 in Appendix G to obtain the dwell times according to the procedure described in Section 5.

Next, we discuss how to extract the order features of a transaction path from the data. There are three different types of order features. The first corresponds to the consumer's demographic information specified in Table 10. The second corresponds to the two attributes in Table 9 of the SKUs in the consumer's past consecutive orders. The third is the *position* of an order in the transaction path. To remove the effect of the extreme value of the *position* on the estimation problem (such as to avoid biased estimation or over-fitting), we adopt a clamp transformation method from the machine learning literature (Kelleher et al. 2015) to cap the *position* such that if an order has a position > 4, we reset its position to 4.

Note that except *attribute1* and *attribute2* in Table 9, all the order features above are categorical variables. Different from numerical variables, the value of each categorical variable provides a convenient label for the category and cannot be directly used in a regression model. We transform each categorical variable to a numeric counterpart as follows. Based on the literature (Singh et al. 2009, Tang et al. 2014), we use dummy coding to transform each categorical variable to a series of variables that equal 0 or 1. For example, there are four categories for the feature *education*: 1 for less than high school, 2 for high school diploma or equivalent, 3 for Bachelor's degree, and 4 for post-graduate degree. The dummy coding introduces three dummy variables denoted as $\mathcal{I}(education = i)$, $i = 2, 3, 4$, where $\mathcal{I}(education = i)$ equals 1 if the education level is i, and 0 otherwise. Each category corresponds to a unique vector of dichotomous variables $(\mathcal{I}(education = 2), \mathcal{I}(education = 3), \mathcal{I}(education = 4))$ shown in Table 1, where "less than high school" coded as $(0, 0, 0)$ is a *reference category*. The regression coefficient of each dummy variable represents the difference between the mean of the dependent variable for the corresponding category and that for the reference category.

Table 1 Dummy coding for the categorical variable "education"

Education	$(\mathcal{I}(education=2), \mathcal{I}(education=3), \mathcal{I}(education=4))$
less than high school	(0, 0, 0)
high school diploma or equivalent	(1, 0, 0)
Bachelor's degree	(0, 1, 0)
post-graduate degree	(0, 0, 1)

We use the numerical variables *attribute1* and *attribute2* as the product attributes of each order. Thus, the dimension of each $\mathbf{x}_{l,k}$, $k=1,\dots,l$, is $d_1=2$, and their coefficients $\boldsymbol{\eta}_1=(\eta_{1,1},\eta_{1,2})$. However, there are 27.10% of the orders requesting SKUs with missing values for the two attributes. Shen et al. (2019) summarize the reasons as follows: "(a) The third-party merchants did not provide the attribute value, especially for certain slow-moving items, or (b) a certain attribute was not applicable to certain SKUs." Therefore, we add two attributes *attribute1_missing* and *attribute2_missing* as categorical variables to indicate whether the values of *attribute1* and *attribute2*, respectively, are missing. See the first column of Table 4 in Section 7.2 for the extracted features.

With the extracted transaction paths $\{path^k\}_{k\in K}$, we estimate the coefficients $\Theta=(a,b,\boldsymbol{\eta},\boldsymbol{\beta})$ of the choice network N by solving the MLE model (23) and predict the choice probabilities. This connects the extracted order features with the transition probabilities of the set of single-dimensional MDPs (10). Based on the estimated coefficients and probabilities, we can compute the personalized threshold and its piecewise-linear approximation in Theorems 2 and 3 respectively.

7. Numerical Experiments

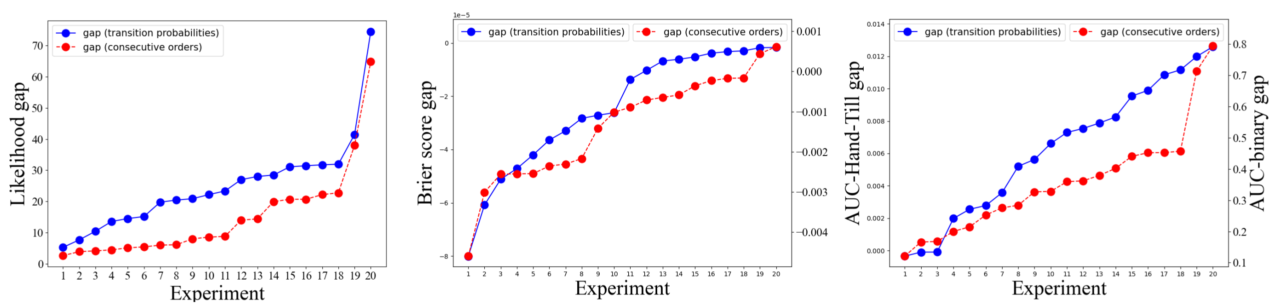
We conduct numerical experiments based on the data processed in Section 6. Section 7.1 compares MLE with two widely-used methods in the statistics and machine learning literature. Section 7.2 computes the personalized threshold and draws some insights. Finally, Section 7.3 benchmarks the personalized threshold-type policy against two heuristics commonly used in the literature.

7.1. Assessing MLE's Accuracy

We solve the MLE model (23) to estimate $\Theta=(a,b,\boldsymbol{\eta},\boldsymbol{\beta})$ and predict (i) the transition probabilities of the single-dimensional MDPs (10) and (ii) the consumer's probability of placing consecutive orders. We compare MLE's accuracy with two estimation methods in the statistics and machine learning literature: kernel conditional density estimate (KCDE) and random forest (RF).

7.1.1. Comparing with KCDE KCDE is a widely used non-parametric method for estimating probability density functions. Appendix B.1 estimates the transition probabilities of the single-dimensional MDPs (10) and the probability of placing consecutive orders using KCDE in class "npcdensbw" of R package "np". Since KCDE is computationally intensive (it runs for one week without any results for 300,000 orders), we randomly sample 500 transaction paths for each experiment to reduce the computation time. We compare MLE and KCDE (with 5-fold cross-validation) for each experiment. We introduce an additional regularization term $-\gamma\|\boldsymbol{\eta}\|$, $\gamma \in \{0.1 \times w | w \in \langle 9 \rangle\} \cup [10]$, in the MLE model (23) to mitigate the over-fitting effect.

We compare MLE and KCDE over 20 experiments based on three performance metrics: likelihood, Brier score, and area under the curve (AUC) (see Appendix C for their definitions). Brier score is a widely-used metric in evaluating probabilistic classifiers with a smaller score implying a better prediction accuracy (Pedregosa et al. 2011). AUC is a popular classification performance metric to assess binary classifiers (denoted as AUC-binary) or multi-class classifiers (denoted as AUC-Hand-Till) (Hand and Till 2001). The larger the AUC, the better the prediction accuracy.



(a) (b) (c)

Figure 5 Comparing MLE's and KCDE's prediction accuracy based on different performance metrics

Figure 5a shows the gap between MLE's likelihood and KCDE's likelihood (MLE's likelihood minus KCDE's likelihood) for predicting the transition probabilities (solid line) and the probability of placing consecutive orders (dashed line). Figure 5b shows the gap between MLE's Brier score and KCDE's Brier score for predicting the transition probabilities (solid line) and the probability of placing consecutive orders (dashed line). Figure 5c shows the gap between MLE's AUC-Hand-Till and KCDE's AUC-Hand-Till for predicting the transition probabilities (solid line) as well as the gap between MLE's AUC-binary and KCDE's AUC-binary for predicting the probability of placing consecutive orders (dashed line). These figures suggest that MLE consistently outperforms KCDE for predicting the transition probabilities and the probability of placing consecutive orders based on all the three performance metrics. In addition, MLE requires much less computation time (total 40 seconds for 20 experiments) than KCDE (total 55.04 hours for 20 experiments).

7.1.2. Comparing with RF We further compare MLE with RF, one of the most powerful probabilistic classifiers in the machine learning literature (Niculescu-Mizil and Caruana 2005). See Appendix B.2 for the details of predicting the probabilities using RF. Note that a direct output of MLE and RF is the prediction of the choice probability of each alternative. We first compare their prediction accuracy for the choice probabilities in the consumer's sequential decision process. Table 2 shows that their performance is very similar in terms of Brier score and AUC-Hand-Till. MLE achieves a slightly better performance in AUC-Hand-Till, but is slightly worse in Brier score. Figure 6 further compares the predicted transition probabilities by the two methods. Define $GAP_l^k(w) = p_{MLE}((\mathbf{x}_l^k, w), (\mathbf{x}_l^k, w+1)) - p_{RF}((\mathbf{x}_l^k, w), (\mathbf{x}_l^k, w+1))$, where $p_{MLE}(\cdot,\cdot)$ and $p_{RF}(\cdot,\cdot)$ represent the transition probabilities predicted by MLE and RF respectively. The gap for each dwell time w is close to 0, indicating that the predictions by the two methods are very close.

Table 2 Comparing MLE and RF in predicting choice probabilities

Metrics	RF	MLE
Brier score	0.029434159	0.030050049
AUC-Hand-Till	0.599577933	0.604655914

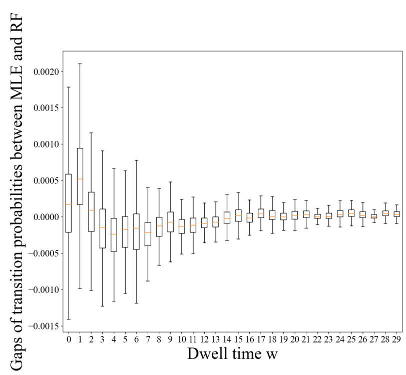
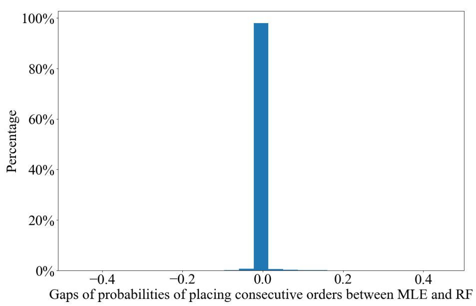


Figure 6 Comparing MLE and RF in predicting transition probabilities

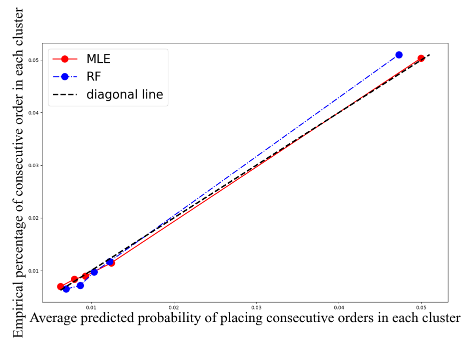
Table 3 compares the accuracy of the two methods in predicting the probability of placing consecutive orders in terms of Brier score and AUC-binary. Both methods give very similar prediction accuracies. Figure 7a shows the distribution of the gaps between the predicted probabilities by the two methods. The figure shows that most gaps are zero, suggesting that their predictions are extremely close. Following the machine learning literature on comparing binary probabilistic classifiers (Niculescu-Mizil and Caruana 2005), we also investigate the *calibration curves* under the two methods. See Appendix C for the procedure of constructing a calibration curve. The nearer the calibration curve is to the diagonal line, the better the prediction is. Figure 7b shows that the calibration curve of MLE is nearer to the diagonal line compared to that of RF, suggesting that MLE slightly outperforms RF in predicting the probability of placing consecutive orders.

Table 3 **Brier score and AUC-binary of MLE and RF for predicting the probability of placing consecutive orders**

Metrics	RF	MLE
Brier score	0.013107706	0.013282569
AUC-binary	0.762402023	0.752302149



(a)



(b)

Figure 7 **Comparing MLE and RF in predicting the probability of placing consecutive orders**

In summary, the predictions of MLE and RF are very close to each other. MLE performs slightly better in terms of AUC-Hand-Till and the calibration plot, but slightly worse in Brier score and AUC-binary than RF. We choose MLE in this paper because of its advantages below.

(a) **Interpretability**: The proposed MLM (13) is based on the random utility maximization framework (12). It models how the consumer chooses among the three alternatives: search and see, leave the system, or place another order. It provides intuitions of the cause of predictions and consistently predicts model results. In contrast, the nonparametric RF method is more like a black box for predictions and its prediction performance heavily depends on the quantity and quality of data.

(b) **Analytical results**: MLM (13) estimated by MLE leads to several nice theoretical results. First, the closed-form expression of the personalized threshold in Theorem 2 can better illustrate its dependence on order features. Second, Corollary 1 provides managers a guideline to select products with desired attributes (when recommending products) in order to control the probability of placing consecutive orders. Finally, the piecewise-linear function with an approximation guarantee in Theorem 3 allows us to visualize the marginal effect of each feature on the threshold and hence identify the important features. In contrast, these results are not available for RF because of its complicated model structure.

(c) **Computational efficiency**: MLE shows a significant advantage over RF in computational efficiency as RF takes 3.87 hours for the hyper-parameter selection, while MLE takes only 0.24 hours in our numerical experiments.

We also compare MLE against KCDE and RF based on their out-of-sample costs in Appendix D. Again, the performance of MLE is better than that of KCDE, but is similar to that of RF.

7.2. Understanding The Personalized Threshold

In this section, we first compute the optimal personalized threshold using the data described in Section 6. Since a large proportion of consumers places a consecutive order within 30 minutes in Figure 1, we set $T = 30$. We normalize the order arrangement cost c to 1. Note that if h is too large, then the retailer will send the pending combined order to the 3PL soon, leading to a large total order arrangement cost. However, if h is too small, then the retailer will hold the pending combined order for a long time, increasing the total holding cost. Thus, the holding cost per period h should be chosen such that the retailer sends the pending combined order to the 3PL as soon as possible, while guaranteeing a pre-specified target total order arrangement cost. For the following numerical experiments, we consider $h = 0.262 \times 10^{-3}, 0.861 \times 10^{-3}, 1.775 \times 10^{-3}$, and 2.788×10^{-3} based on different target total order arrangement costs. See Appendix F for a procedure of choosing these values of h. Figure 8 presents the distribution of the optimal thresholds based on Theorem 2 for all the orders in the data set. Most of the thresholds computed are no more than 10 periods.

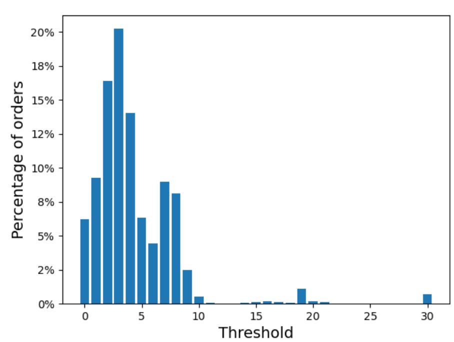


Figure 8 Distribution of the personalized thresholds

Next, we apply the piecewise-linear approximation in Theorem 3 to investigate the marginal effect of each order feature on the personalized threshold. We obtain relevant managerial insights by studying these marginal effects based on the data described in Section 6.

7.2.1. The Accuracy of The Piecewise-Linear Approximation

The last three rows of Table 4 present three metrics to assess the piecewise-linear approximation's accuracy. The first is the analytical *bound* in Theorem 3 of the gap between the optimal threshold and its approximation.

The bound equals 1 for all the different h values, suggesting that the error is at most one period. The second is an *empirical gap* between the optimal threshold in Theorem 2 and its piecewise-linear approximation in Theorem 3 for each order. The empirical gap is either 0 or 1, which is consistent with the analytical bound derived earlier. The first and second entries in each bracket represent the numbers of orders with an empirical gap equals 0 and 1 respectively. The results show that 99.91% to 99.98% of the orders have an empirical gap equal to 0. The last metric is a *cost gap* $GAP_0 = (C_{pl} - C_p)/C_p$, where C_p and C_{pl} are the average total costs incurred under the optimal threshold and its approximation respectively. Table 4 shows that the resulting cost gap GAP_0 based on the data is negligible. In summary, the piecewise-linear function in Theorem 3 approximates the optimal threshold extremely well.

Table 4 The effects of order features on the threshold and the piecewise-linear approximation's accuracy

h	0.262×10^{-3}	0.861×10^{-3}	1.775×10^{-3}	2.788×10^{-3}
$\mathcal{I}(plus = 1)$	1.07	1.07	1.07	1.07
$\mathcal{I}(gender = Male)$	-0.77	-0.77	-0.77	-0.77
$\mathcal{I}(age \leq 15)$	-5.67	-5.67	-5.67	-5.67
$\mathcal{I}(age = 26-35)$	-1.96	-1.96	-1.96	-1.96
$\mathcal{I}(age = 36-45)$	-2.26	-2.26	-2.26	-2.26
$\mathcal{I}(age = 46-55)$	-0.73	-0.73	-0.73	-0.73
$\mathcal{I}(age \geq 56)$	-1.07	-1.07	-1.07	-1.07
$\mathcal{I}(marital\_status = Single)$	-0.66	-0.66	-0.66	-0.66
$\mathcal{I}(education = high\ school)$	-0.15	-0.15	-0.15	-0.15
$\mathcal{I}(education = Bachelor's\ degree)$	0.91	0.91	0.91	0.91
$\mathcal{I}(education = post\text{-}graduate\ degree)$	0.59	0.59	0.59	0.59
$\mathcal{I}(purchase\_power = 2)$	1.18	1.18	1.18	1.18
$\mathcal{I}(purchase\_power = 3)$	1.70	1.70	1.70	1.70
$\mathcal{I}(purchase\_power = 4)$	-0.38	-0.38	-0.38	-0.38
$\mathcal{I}(city\_level = 2)$	0.26	0.26	0.26	0.26
$\mathcal{I}(city\_level = 3)$	0.50	0.50	0.50	0.50
$\mathcal{I}(city\_level = 4)$	0.53	0.53	0.53	0.53
$\mathcal{I}(city\_level = 5)$	-1.82	-1.82	-1.82	-1.82
$\mathcal{I}(user\_level = 0)$	-0.80	-0.80	-0.80	-0.80
$\mathcal{I}(user\_level = 2)$	1.56	1.56	1.56	1.56
$\mathcal{I}(user\_level = 3)$	2.55	2.55	2.55	2.55
$\mathcal{I}(user\_level = 4)$	6.36	6.36	6.36	6.36
$\mathcal{I}(enterprise = 1)$	18.35	18.35	18.35	18.35
$\mathcal{I}(position = 2)$	12.54	12.54	12.54	12.54
$\mathcal{I}(position = 3)$	22.91	22.91	22.91	22.91
$\mathcal{I}(position = 4)$	29.78	29.78	29.78	29.78
$attribute1$	15.65	15.65	15.65	15.65
$\mathcal{I}(attribute1\_missing = 1)$	-17.37	-17.37	-17.37	-17.37
$attribute2$	16.05	16.05	16.05	16.05
$\mathcal{I}(attribute2\_missing = 1)$	34.03	34.03	34.03	34.03
intercept	7.80	0.03	-4.73	-7.72
bound	1	1	1	1
empirical gaps	(301,542; 269)	(301,709; 102)	(301,589; 222)	(301,754; 57)
cost gap GAP_0	-2.34×10^{-7}	-2.76×10^{-7}	-1.24×10^{-6}	-3.07×10^{-7}

Table 4 also shows the coefficients of the piecewise-linear function in Theorem 3. Based on Corollary 1, if an order feature is a numerical variable, its coefficient represents its marginal effect on the threshold. In contrast, if a feature is a categorical variable, the coefficient of each dummy variable represents the difference between the mean threshold for the corresponding category and that for the reference category (omitted in Table 4). It is worth noting that according to (20), the cost parameter h only affects the intercept of $f_l(\mathbf{x};\Theta)$. The marginal effect of each order feature is not affected by varying h. Thus, for each feature we observe the same coefficient for different values of h in Table 4. We describe some useful managerial insights from Table 4 regarding how the order features affect the threshold as follows.

7.2.2. Managerial Insights from The Piecewise-Linear Approximation

(i) *The personalized threshold-type policy suggests a significantly higher threshold for enterprise users and for orders with a higher position.* The coefficient of $\mathcal{I}(enterprise = 1)$ is 18.35, implying that the threshold for an *enterprise* user is at least 18 minutes longer than that for an individual user (reference category) who shares all other order features with the enterprise user. Furthermore, the second last column of Table 13 in Appendix G shows that an enterprise user is much more likely to make consecutive orders than an individual user. This suggests a positive correlation between the probability of placing consecutive orders and the holding threshold, which is consistent with Corollary 1. We have similar observations on the feature *position*. Table 4 shows that a significantly higher threshold should be used for $position = 2, 3$, or 4 compared to $position = 1$ (reference category). Table 14 in Appendix G shows that $position = 2, 3$, or 4 has a significantly higher empirical percentage of orders having a consecutive order compared to $position = 1$.

(ii) *The personalized threshold-type policy suggests a higher threshold for plus users.* The coefficient of $\mathcal{I}(plus = 1)$ in Table 4 is positive, suggesting a higher threshold for a plus user compared to a non-plus user (reference category) who shares all other order features with the plus user. The plus users pay a smaller fee for placing a new order, incentivizing them to place small orders frequently instead of a single large order. Our policy sets a higher threshold for them.

(iii) *The threshold for female users (reference category) is higher than that for male users who shares all other order features with the female users.* This observation aligns with our intuition that the female users generally tend to spend a longer time in online shopping and are more likely to be persuaded by online recommendations.

(iv) *Users in the age group of 16-25 (reference category) should be assigned a higher threshold than any other age groups that share all other order features with the reference category.* This reflects the age range of the main consumer group shopping online.

(v) *The threshold for tier-1 cities (reference category) is lower than that for tier-2, 3, and 4 cities but higher than that for tier-5 cities that share all other order features with the reference category.* Since online shopping requires a certain buying power and consumes time, the above observation suggests two other factors that may affect online shopping behavior: income level and consumers' pace of life. Consumers in tier-1 cities have a high income level but a fast pace of life. They tend to spend a shorter time searching while shopping online.

7.3. The Value of Personalizing Order-holding Decisions

Here, we assess the value of using the personalized threshold-type policy by comparing it with two widely-used benchmarks in the literature. The first benchmark is a *myopic policy* in the order-fulfillment literature (Xu et al. 2009, Acimovic and Graves 2015, Jasin and Sinha 2015, Wei et al. 2017), which was reported to be implemented in Amazon.com (Ng 2012). Under the myopic policy, any received orders will be sent to the 3PL immediately. In other words, the holding threshold is zero for all the orders. The second benchmark is a *static policy*, which adopts a constant threshold for all the orders regardless of their features. Under the static policy, any pending combined order is held if the dwell time is less than the static threshold, but sent to the 3PL otherwise. See Appendix E for finding an optimal constant threshold for the static policy.

Let C_m, C_s, and C_p denote the total costs under the myopic, static, and personalized threshold-type policies respectively. Following the machine learning literature, we use 10-fold cross-validation to compare the personalized threshold-type policy with the two benchmarks. Specifically, we divide the transaction paths $\{path^k\}_{k\in[K]}$ into 10 even groups. We use one group for testing and the remaining groups for training. We repeat this process 10 times, each with a different testing group. We sum up the 10 testing costs in the cross-validation to obtain the total cost under each policy. The second column of Table 5 shows $(C_m; C_s; C_p)$. Define the cost gap between a benchmark and the personalized threshold-type policy as $CGAP_i = (C_i - C_p)/C_p \times 100\%$, where $i \in \{m, s\}$. Let A_m, A_s, and A_p denote the total order arrangement costs under the myopic, static, and personalized threshold-type policies respectively. Define the gap between the total order arrangement costs of a benchmark and the personalized threshold-type policy as $AGAP_i = (A_i - A_p)/A_p \times 100\%$, where $i \in \{m, s\}$. The third and fourth columns of Table 5 show $(CGAP_m, CGAP_s)$ and $(AGAP_m, AGAP_s)$ respectively. The first, second, and third entries of the last column represent the average holding times under the myopic, static, and personalized threshold-type policies respectively.

Table 5 Performance of the personalized threshold-type policy and benchmarks

$h(\times 10^{-3})$	**Total costs**	$(CGAP_m, CGAP_s)$(%)	$(AGAP_m, AGAP_s)$ (%)	**Holding times**
0.262	(301,811; 298,538; 298,152)	(1.227, 0.129)	(1.545, 0.073)	(0.0, 14.35, 12.66)
0.861	(301,811; 299,907; 299,288)	(0.843, 0.207)	(1.21, 0.095)	(0.0, 5.56, 4.91)
1.775	(301,811; 300,853; 299,786)	(0.675, 0.356)	(0.86, 0.076)	(0.0, 2.62, 1.68)
2.788	(301,811; 301,487; 300,085)	(0.575, 0.467)	(0.675, 0.091)	(0.0, 1.72, 0.97)

Table 6 shows the reductions in the total cost and in the total order arrangement cost by the personalized threshold-type policy for different *user levels*. The first and second entries of each tuple represent the reductions compared to the myopic and static policies respectively. The table

Table 6 Reductions in the total cost, total order arrangement cost, and holding time for each user level

	$h(\times 10^{-3})$	*user_level*=1	*user_level*=2	*user_level*=3	*user_level*=4	*enterprise*
Total cost (%)	0.262	(0.28, 0.05)	(0.42, 0.05)	(0.61, 0.03)	(1.95, 0.02)	(36.41, 7.91)
	0.861	(0.03, 0.09)	(0.12, 0.06)	(0.29, 0.08)	(1.46, -0.01)	(31.49, 13.12)
	1.775	(0.01, 0.23)	(0.03, 0.15)	(0.1, 0.09)	(1.25, 0.16)	(27.71, 17.2)
	2.788	(0.01, 0.34)	(0.03, 0.26)	(0.09, 0.24)	(0.97, 0.19)	(24.76, 17.72)
Total order arrangement cost (%)˙	0.262	(0.52, -0.08)	(0.7, -0.04)	(0.91, -0.03)	(2.37, 0.06)	(37.5, 8.33)
	0.861	(0.16, -0.25)	(0.35, -0.18)	(0.61, -0.08)	(2.14, 0.18)	(34.08, 14.77)
	1.775	(0.02, -0.22)	(0.05, -0.29)	(0.13, -0.33)	(1.78, 0.21)	(31.8, 20.38)
	2.788	(0.02, -0.12)	(0.05, -0.19)	(0.11, -0.21)	(1.11, -0.15)	(29.95, 22.02)
Holding time (%)	0.262	52.8	32.78	21.57	-10.49	-54.16
	0.861	263.22	95.48	44.36	-31.85	-86.59
	1.775	3323.89	1938.68	930.87	-22.52	-93.93
	2.788	3112.07	1924.4	1162.44	94.61	-95.85

also shows the reduction in the average holding time by the personalized threshold-type policy compared to the static policy. We have the following observations from these results.

1. *Compared to the myopic policy, the personalized threshold-type policy reduces the total cost by having fewer order arrangements.* Table 5 shows that the personalized threshold-type policy reduces the total cost by 0.575% to 1.227% compared to the myopic policy (see $CGAP_m$). The cost reduction is due to fewer order arrangements (see $AGAP_m$). For example, for $h = 0.262 \times 10^{-3}$ the personalized threshold-type policy reduces the total order arrangement cost by 1.545% compared to the myopic policy. This cost reduction by consolidating the orders has a significant impact in practice because of the large number of consumer orders in online retailing.

2. *Compared to the static policy, personalizing the threshold reduces the total cost by having fewer order arrangements and shorter holding times.* According to $CGAP_s$ in Table 5, the personalized threshold-type policy reduces the total cost by 0.129% to 0.467% compared to the static policy. This demonstrates the value of personalizing the threshold by considering the features of the orders placed by a consumer, which leads to fewer order arrangements and shorter holding times (see columns 4 and 5 of Table 5).

3. *Optimizing order-holding decisions is significantly more valuable for enterprise users.* Table 6 shows that the personalized threshold-type policy reduces the total cost more significantly for the enterprise users. Among the enterprise users, it reduces the total cost of the myopic policy by 24.76% to 36.41% and the total cost of the static policy by 7.91% to 17.72%. Interestingly, the cost reduction is caused by having fewer order arrangements and *longer* holding times. The significantly larger percent reduction in the total cost compared to other user levels is partly due to the high percentage of consecutive orders made by the enterprise users. Thus, optimizing the order-holding decisions is significantly more valuable for the enterprise users, suggesting that personalizing the threshold becomes even more crucial as the proportion of the enterprise users increases. This insight

is quite implementable in practice because it is straightforward to identify the enterprise users when they are labeled during registration.

8. Conclusion

A significant percentage of consumers place consecutive, separate orders within a short duration when shopping online (see Figure 1, which is based on JD.com's data). An online retailer can hold any order(s) of each consumer for a limited time interval to reduce the total order arrangement cost. This approach is also known as order consolidation for a *single* consumer. However, we do not want to hold the pending order(s) for too long so that we can maintain a responsive fulfillment process for the consumer. We analyze a relevant trade-off in consolidating consecutive, separate orders placed by the same consumer in online retailing that has not been studied in the literature.

We model the online retailer's order-holding problem for a single consumer as an MDP. Due to the curse of dimensionality, we do not solve this MDP directly. Instead, we first construct a set of single-dimensional MDPs. We show that the optimal policy of the original MDP can be characterized by *personalized* thresholds that can be obtained by solving the set of single-dimensional MDPs (see Theorem 1), which are much more tractable. The holding threshold for each pending set of orders depends on the transition probabilities of an associated single-dimensional MDP.

To estimate the transition probabilities, we apply a consumer's sequential decision model to characterize the consumer choice behavior after placing an order. This allows us to construct a closed-form relation (16) between the transition probabilities and the order features. Based on (16), we personalize the optimal threshold by analytically expressing it in terms of the order features (see Theorem 2). To better understand the relation between the optimal personalized threshold and the order features, we propose a piecewise-linear approximation to the threshold with an upper bound on the approximation gap (see Theorem 3). We also show that the optimal personalized threshold is positively correlated with the consumer's probability of placing consecutive orders. We analytically express this probability in order features (see Proposition 2), which helps the retailer predict the consumer's probability of placing a consecutive order after receiving an order.

We apply the MLE method to estimate the coefficients of the consumer's sequential decision model using the data of 2020 MSOM Data Driven Research Challenge. Our extensive numerical tests suggest that the MLE method accurately predicts the transition probabilities of each single-dimensional MDP and the probability of placing consecutive orders by a consumer compared to the KCDE and RF methods that are widely used in the literature.

Our numerical studies suggest that the piecewise-linear function in Theorem 3 serves as a good approximation to the optimal threshold in Theorem 2. By using the piecewise-linear approximation, we can assess the marginal effect of each order feature on the personalized threshold. Based on the data of 2020 MSOM Data Driven Research Challenge, we obtain the following interesting and relevant managerial insights from the coefficients of the piecewise-linear function in Table 4: (i) *A significantly higher threshold should be used for enterprise users and for orders with a higher position.* Furthermore, both enterprise users and orders with a higher position have a higher probability of placing consecutive orders, implying a positive correlation between this probability and the holding threshold, which is consistent with Corollary 1. (ii) *A higher threshold should be used for plus users.* (iii) *The threshold for female users is higher than that for male users.* (iv) *Users in the age group of 16-25 have a higher threshold than other age groups.* (v) *The threshold for tier-1 cities is lower than that for tier-2, 3, and 4 cities but higher than that for tier-5 cities.*

Comparing the personalized threshold-type policy with the myopic and the static policies, we observe the following: (a) *Compared to the myopic policy, personalizing the threshold reduces the total cost by having fewer order arrangements.* (b) *Compared to the static policy, personalizing the threshold reduces the total cost by having fewer order arrangements and shorter holding times.* (c) *Optimizing order-holding decisions is significantly more valuable for enterprise users.* Finally, although our work applies to consumer-facing consolidation operations in e-commerce, the proposed feature-based data analytics approach can be generalized to other personalized policies.

Acknowledgement

The authors thank the Department Editor, Associate Editor, and two anonymous referees for their valuable comments that have substantially improved the paper. Portions of this paper were presented in University of Chinese Academy of Sciences, 2022 and the 13th POMS-HK International Conference, Hong Kong, 2023. The authors thank the audiences for many insightful comments and stimulating questions. This work was done while S. Chen was a research fellow at Nanyang Technological University. S. Chen acknowledges the financial support from the National Natural Science Foundation of China [Grants 72201237 and 72231009]. Z. Yan was partly supported by Singapore Ministry of Education Academic Research Fund Tier 1 [Grant RG17/21] and Tier 2 [Grant MOE2019-T2-1-045] as well as Neptune Orient Lines Fellowship [Grant NOL21RP04]. Y. F. Lim is grateful for the support from the Lee Kong Chian School of Business, Singapore Management University [Maritime and Port Authority Research Fellowship], the Association of South-East Asian Nations Business Research Initiative Grant [Grant G17C20421], the Research Grants Council of Hong Kong [Grants 15501920 and 15501221], and the Key Program of National Natural Science Foundation of China [No. 71931009].

References

Acimovic J, Graves SC (2015) Making better fulfillment decisions on the fly in an online retail environment. *Manufacturing & Service Operations Management* 17(1):34-51.

Ahipaşaoğlu SD, Arikan U, Natarajan K (2019) Distributionally robust Markovian traffic equilibrium. *Transportation Science* 53(6):1546-1562.

Alptekinoğlu A, Semple JH (2016) The exponomial choice model: A new alternative for assortment and price optimization. *Operations Research* 64(1):79-93.

Akamatsu T (1996) Cyclic flows, Markov process and stochastic traffic assignment. *Transportation Research Part B: Methodological*, 30(5):369-386.

Akamatsu T (1997) Decomposition of path choice entropy in general transport networks. *Transportation Science*, 31(4):349-362.

Anderson TW, Goodman LA (1957) Statistical inference about Markov chains. *The Annals of Mathematical Statistics*, 89-110.

Andrews JM, Farias VF, Khojandi AI, Yan CM (2019) Primal-Dual Algorithms for Order Fulfillment at Urban Outfitters, Inc. *INFORMS Journal on Applied Analytics* 49(5):355-370.

Bashtannyk DM, Hyndman RJ (2001) Bandwidth selection for kernel conditional density estimation. *Computational Statistics & Data Analysis*, 36(3):279-298.

Bell MG (1995) Alternatives to Dial's logit assignment algorithm. *Transportation Research Part B: Methodological* 29(4):287-295.

Blanchet J, Gallego G, Goyal V (2016) A Markov chain approximation to choice modeling. *Operations Research* 64(4):886-905.

Bowman AW (1980) A note on consistency of the kernel method for the analysis of categorical data. *Biometrika* 67(3):682-684.

Business Statistics (2021) JD Statistics, User Counts, and Facts. https://expandedramblings.com/index.php/by-the-numbers-15-amazing-jd-com-stats/. Accessed Jul 16, 2021.

Çeven E, Gue KR (2017) Optimal wave release times for order fulfillment systems with deadlines. *Transportation Science* 51(1):52-66.

Collins L (1974) Estimating Markov Transition Probabilities from Micro Unit Data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 23(3):355-371.

De Koster R, Le-Duc T, Roodbergen KJ (2007) Design and control of warehouse order picking: A literature review. *European Journal of Operational Research* 82(2):481-501.

Ed Smith (2020) The Increasing Demand of Consumer Delivery Expectations. *Retail Technology Review*. https://www.retailtechnologyreview.com/articles/2020/02/25/the-increasing-demand-of-consumer-delivery-expectations/.

eMarketer (2021) Global Ecommerce Update 2021: Worldwide Ecommerce Will Approach $5 Trillion This Year. https://www.emarketer.com/. Accessed Feb 15, 2021.

Flores A, Berbeglia G, Van Hentenryck P (2019) Assortment optimization under the sequential multinomial logit model. *European Journal of Operational Research* 273(3):1052-1064.

Gallego G, Li A, Truong VA, Wang X (2020) Approximation algorithms for product framing and pricing. *Operations Research* 68(1):134-160.

Gzara F, Samir E, Yildiz U, Baloch G (2020) Data-Driven Modeling and Optimization of the Order Consolidation Problem in E-Warehousing. *INFORMS Journal on Optimization* 2(4):273-296.

Hall P, Racine J, Li Q (2004) Cross-validation and the estimation of conditional probability densities. *Journal of the American Statistical Association* 99(468):1015-1026.

Hand DJ, Till RJ (2001) A simple generalisation of the area under the ROC curve for multiple class classification problems. *Machine learning* 45(2):171-186.

Jasin S, Sinha A (2015) An LP-based correlated rounding scheme for multi-item ecommerce order fulfillment. *Operations Research* 63(6):1336-1351.

Kelleher JD, Mac Namee B, D'arcy A (2015) Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies. *MIT Press*.

Lei Y, Jasin S, Sinha A (2018) Joint dynamic pricing and order fulfillment for e-commerce retailers. *Manufacturing & Service Operations Management* 20(2):269-284.

Li Q, Racine JS (2007) Nonparametric econometrics: theory and practice. *Princeton University Press*.

Lim YF, Jiu S, Ang M (2021) Integrating anticipative replenishment allocation with reactive fulfillment for online retailing using robust optimization. *Manufacturing & Service Operations Management* 23(6):1616-1633.

Liu N, Ma Y, Topaloglu H (2020) Assortment optimization under the multinomial logit model with sequential offerings. *INFORMS Journal on Computing* 32(3):835-853.

Luce RD. 1959. Individual Choice Behavior. *John Wiley & Sons*.

Macrotrends.net. Amazon profit margin 2006–2018. https://www.macrotrends.net/stocks/charts/AMZN/amazon/profit-margins. Accessed July 8, 2019.

Niculescu-Mizil A, Caruana R (2005) Predicting good probabilities with supervised learning. In *Proceedings of the 22nd International Conference on Machine Learning*, 625-632.

Ng AY (2004) Feature selection, L 1 vs. L 2 regularization, and rotational invariance. In *Proceedings of the Twenty-First International Conference on Machine Learning*, 78.

Ng CK (2012) Inbound supply chain optimization and process improvement. Doctoral dissertation, Massachusetts Institute of Technology.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, ... Duchesnay E (2011) Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12:2825-2830.

Richter F (2019) Amazon's escalating logistics costs. *Statista* **(March 1)**, https://www.statista.com/chart/17207/amazon-shipping-and-fulllment-costs/.

Rosenblatt M (1969) Conditional probability density and regression estimators. *Multivariate Analysis II* 25:31.

Roussas GG (1969) Nonparametric estimation of the transition distribution function of a Markov process. *The Annals of Mathematical Statistics* 40(4):1386-1400.

Rouwenhorst B, Reuter B, Stockrahm V, van Houtum GJ, Mantel RJ, Zijm WH (2000) Warehouse design and control: Framework and literature review. *European Journal of Operational Research* 122(3):515-533.

Shen M, Tang CS, Wu D, Yuan R, Zhou W (2020) JD.com: Transaction Level Data for the 2020 MSOM Data Driven Research Challenge. *Manufacturing & Service Operations Management*. https://doi.org/10.1287/msom.2020.0900.

Singh S, Kubica J, Larsen S, Sorokina D (2009) Parallel large scale feature selection for logistic regression. In *Proceedings of the 2009 SIAM International Conference on Data Mining*, 1172-1183. Society for Industrial and Applied Mathematics.

Stenius O, Karaarslan AG, Marklund J, De Kok AG (2016) Exact analysis of divergent inventory systems with time-based shipment consolidation and compound Poisson demand. *Operations Research* 64(4):906-921.

Tang J, Alelyani S, Liu H (2014) Feature selection for classification: A review. *Data Classification: Algorithms and Applications*, 37.

Train KE (2009) Discrete choice methods with simulation. *Cambridge University Press*.

Wei L, Jasin S, Kapuscinski R (2017) Shipping Consolidation with Delivery Deadline and Expedited Shipment Options. Preprint, submitted February 22, https://ssrn.com/abstract=2920899.

Williams HC (1977) On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A* 9(3):285-344.

Xu PJ, Allgor R, Graves SC (2009) Benefits of reevaluating real-time order fulfillment decisions. *Manufacturing & Service Operations Management* 11(2):340-355.

Yakowitz SJ (1985) Nonparametric density estimation, prediction, and regression for Markov sequences. *Journal of the American Statistical Association* 80(389):215-221.

Yan Z (2020) Pricing with consumer sequential choice model and its application in the design of a threshold-type promotion. Preprint, submitted July 1, https://ssrn.com/abstract=3623031.

Yan Z, Cheng C, Natarajan K, Teo CP (2022) A representative consumer model in data-driven multi-product pricing optimization. *Management Science* 68(8), 5557-6354.

Online Appendix to
"Managing The Personalized Order-Holding Problem in Online Retailing"

Appendix

A. Proofs

LEMMA 2 (PARTIALLY-HOLDING IS SUBOPTIMAL). *It is always suboptimal to hold only part of consecutive orders in the current pending orders and send others to the 3PL.*

Proof of Lemma 2. Denote π^* as the optimal order-holding policy, then the optimal total expected cost after taking the partially-holding policy in the current period, denoted as V_1, can be obtained by applying the optimal policy π^* in subsequent periods. To show it is suboptimal, it suffices to construct another policy whose total expected cost, denoted as V_2, is less than V_1 which is the least cost that the partially-holding policy can achieve.

Denote the set of orders held and sent according to the partially-holding policy in the current period as B and A respectively. The constructed policy is to hold the whole pending orders in the current period, which only triggers cost at h in the current period compared to $c+h$ for the partially-holding counterpart. The subsequent actions for the constructed policy will hold all orders in A and B until the beginning of period t when some orders, denoted as B', in B will be sent to the 3PL according to the optimal policy π^*. Then, the constructed policy at the beginning of period t is to send all orders in A and B' to the 3PL and hold the rest of orders if there remains some. After period t, the subsequent decisions for the constructed policy follow the optimal counterpart π^*. Due to the cost structure, the subsequent actions constructed above after the current period will lead to the same cost in subsequent periods as the optimal policy π^*, so the elaborated policy, which leads to less cost in the current period, has less cost than partially-holding counterpart which completes this proof. □

Proof of Lemma 1. Firstly, we show that the optimal order-holding decision at states $(\mathbf{x}_l, T, v=1)$ and $(\mathbf{x}_L, 0, v=1)$ must be to send the current pending orders to the 3PL immediately. We only provide proof for the states $(\mathbf{x}_l, T, v=1)$, and the proof for the state $(\mathbf{x}_L, 0, v=1)$ is similar, so ignored. To show this phenomenon is true, recall that the only subsequent state of $(\mathbf{x}_l, w)$ is $(\mathbf{x}_l, w+1)$ for $w \geq T$, so $p((\mathbf{x}_l, w), (\mathbf{x}_l, w+1)) = 1$ (the consumer has left the market). Then, problem (3) for $w \geq T$ will turns to

$$
\begin{aligned}
V(\mathbf{x}_l, w, v=1) &= \min_{z\in\{0,1\}} z\{h + p((\mathbf{x}_l, w), (\mathbf{x}_l, w+1)) \cdot V(\mathbf{x}_l, w+1, v=1)\} && (24)\\
&\quad + (1-z)\{c + p((\mathbf{x}_l, w), (\mathbf{x}_l, w+1)) \cdot V(\mathbf{x}_l, w+1, v=0)\}\\
&= \min_{z\in\{0,1\}} z\{h + V(\mathbf{x}_l, w+1, v=1)\} + (1-z)\cdot c && (25)
\end{aligned}
$$

where the second equality comes from the fact that $p((\mathbf{x}_l, w), (\mathbf{x}_l, w+1)) = 1$ and $V(\mathbf{x}_l, w+1, v=0) = 0$ (since $(\mathbf{x}_l, w+1, v=0)$ indicates that the current pending orders have been sent to the 3PL and no new order will come). Prove by contradiction. If the optimal decision at state $(\mathbf{x}_l, T, v=1)$ is $z^* = 1$, then we have $V(\mathbf{x}_l, T, v=1) = h + V(\mathbf{x}_l, T+1, v=1) < c$ which implies that $V(\mathbf{x}_l, T+1, v=1) < c-h$. By (25), the optimal decision at state $(\mathbf{x}_l, T+1, v=1)$ must be $z^* = 1$ otherwise $V(\mathbf{x}_l, T+1, v=1) = c$ which contradicts $V(\mathbf{x}_l, T+1, v=1) < c-h$. So the optimal decision at state $(\mathbf{x}_l, T+1, v=1)$ must be $z^* = 1$ which implies that $V(\mathbf{x}_l, T+1, v=1) = h + V(\mathbf{x}, T+2, v=1) < c-h$. So we have $V(\mathbf{x}, T+2, v=1) < c-2h$. Follow the same procedures, we will show that $V(\mathbf{x}_l, w, v=1) < c-(w-T)\cdot h$ for $w \geq T$. Here comes the contradiction because $V(\mathbf{x}_l, w, v=1)$ is negative for large enough w. So, the optimal decision at state $(\mathbf{x}_l, T, v=1)$ is $z^* = 0$, that is $V(\mathbf{x}_l, T, v=1) = c$ which is exactly (8). So we have proved that the optimal policy at these states must be to send the current to 3PL immediately, now we turn to characterize their optimal total expected cost.

By the optimal policy at states $(\mathbf{x}_l, T, v=1)$ and $(\mathbf{x}_L, 0, v=1)$, it is easy to check their corresponding optimal total expected cost. Firstly, at state $(\mathbf{x}_l, T, v=0)$, the current pending orders have been sent to the 3PL and no new order will come, so $V(\mathbf{x}_l, T, v=0)=0$. At states $(\mathbf{x}_l, T, v=1)$ and $(\mathbf{x}_L, 0, v=1)$, since the optimal decision is to send the current pending orders to the 3PL immediately, by (25), their corresponding optimal cost must equal to c. So we complete the proof of Lemma 1. □

Proof of Proposition 1. Recall the large scale MDP problem (5) with holding state $v=1$, subtract $V(\mathbf{x}_l, w, v=0)$ from both side of the first equation in (5), we have

$$
\begin{aligned}
& V(\mathbf{x}_l, w, v=1) - V(\mathbf{x}_l, w, v=0) \\
&= \min_{z\in\{0,1\}} z\left\{ h + \sum_{(\mathbf{x}',w')\in\sigma((\mathbf{x}_l,w))} p((\mathbf{x}_l,w),(\mathbf{x}',w')) \begin{pmatrix} V(\mathbf{x}',w',v'=1) \\ -V(\mathbf{x}',w',v'=\mathcal{I}((\mathbf{x}',w')\neq(\mathbf{x}_l,w+1))) \end{pmatrix} \right\} \\
&+ (1-z)\{c\}, \forall w \in \langle T-1\rangle. \\
&= \min_{z\in\{0,1\}} z\{h + p((\mathbf{x}_t,w),(\mathbf{x}_l,w+1))(V(\mathbf{x}_l,w+1,v=1) - V(\mathbf{x}_l,w+1,v=0))\} \\
&+ (1-z)\cdot c.
\end{aligned} \tag{26}
$$

Since subtracting $V(\mathbf{x}_l, w, v=0)$ has no effect on the optimal decision of the original optimization problem in (5), problem (26) must share the same optimal decision as the original problem (5) at state $(\mathbf{x}_l, w, v=1)$. Define $\Delta_{\mathbf{x}_l}(w) = V(\mathbf{x}_l, w, v=1) - V(\mathbf{x}_l, w, v=0)$ (hence $\Delta_{\mathbf{x}_l}(T) = V(\mathbf{x}, T, v=1) - V(\mathbf{x}_l, T, v=0) = c$ by (7) and (8)), then (26) turns to the following single-dimensional MDP problem (27):

$$
\begin{aligned}
& \Delta_{\mathbf{x}_l}(w) = \min_{z\in\{0,1\}} z\{h + p((\mathbf{x}_t,w),(\mathbf{x}_l,w+1))\cdot\Delta_{\mathbf{x}_l}(w+1)\} + (1-z)c, \forall 0\le w\le T-1 \\
& \Delta_{\mathbf{x}_l}(T) = c, \text{ for } w=T
\end{aligned} \tag{27}
$$

whose optimal decision at state w, $z^*_{\mathbf{x}_l}(w)$, coincides with the optimal order-holding decision of the original problem (5) at state $(\mathbf{x}_l, w, v=1)$, $z^*(\mathbf{x}_l, w, v=1)$. This completes the proof of Proposition 1. □

Proof of Theorem 1. By Proposition 1, for each cluster $\mathfrak{E}_{\mathbf{x}_l} = \{(\mathbf{x}_l, w) : 0 \le w \le T\}$, we know that the optimal order-holding policy at state $(\mathbf{x}_l, w, v=1)$ with $(\mathbf{x}_l, w) \in \mathfrak{E}_{\mathbf{x}_l}$ made by the online retailer can be characterized by the single-dimensional MDP problem as follows:

$$
\begin{aligned}
& \Delta_{\mathbf{x}_l}(w) = \min_{z\in\{0,1\}} z\{h + p((\mathbf{x}_l,w),(\mathbf{x}_l,w+1))\cdot\Delta_{\mathbf{x}_l}(w+1)\} + (1-z)c, \forall 0\le w\le T-1 \\
& \Delta_{\mathbf{x}_l}(T) = c, \text{ for } w=T.
\end{aligned} \tag{28}
$$

Firstly, we characterize the optimal decision of the above single-dimensional MDP problem, and then using their connections stated in Proposition 1 to characterize the optimal policy for the original MDP problem (5). Denote $S(\mathbf{x}_l)$ as the first $w \in \langle T-1\rangle$ such that $h + p((\mathbf{x}_l,w),(\mathbf{x}_l,w+1))\cdot \Delta_{\mathbf{x}_l}(w+1) \ge c$, that is $S(\mathbf{x}_l) = \min\{w \in \langle T-1\rangle | h + p((\mathbf{x}_l,w),(\mathbf{x}_l,w+1))\cdot\Delta_{\mathbf{x}_l}(w+1)\ge c\}$ (set to T if the set is empty). Then, for any $w < S(\mathbf{x}_l)$, we must have $h + p((\mathbf{x}_l,w),(\mathbf{x}_l,w+1))\cdot\Delta_{\mathbf{x}_l}(w+1) < c$ which implies that the optimal decision is $z^* = 1$. While for $w = S(\mathbf{x}_l)$, the optimal decision is $z^* = 0$.

With the $S(\mathbf{x}_l)$ defined above characterizing the optimal decision of problem (28), we are ready to characterize the structure of optimal policy for the original MDP problem (5). We prove the structure of the optimal policy by considering the following two cases stated in Theorem 1.

Case 1. If no new order is received for $S(\mathbf{x}_l)$ period. When receiving the order, the current pending orders are being held, so the initial state is $(\mathbf{x}_l, 0, v=1)$. By Proposition 1, the optimal order-holding decision at state $(\mathbf{x}_l, 0, v=1)$ is the same as the counterpart of problem (28) at state 0. So, the optimal order-holding decision at state $(\mathbf{x}, 0, v=1)$ is to hold the current pending orders since $0 < S(\mathbf{x}_l)$. Since no new order is received for $S(\mathbf{x}_l)$ period, the state in the next period is $(\mathbf{x}_l, 1, v=1)$. By the same statement, the pending orders will be held until the beginning of period $S(\mathbf{x}_l)$ in which the state is $(\mathbf{x}_l, S(\mathbf{x}_l), v=1)$. The optimal decision of problem (28) at state $S(\mathbf{x}_l)$ is $z^*=0$, so does the optimal decision for the original MDP problem (5) at state $(\mathbf{x}_l, S(\mathbf{x}_l), v=1)$ by Proposition 1. That is to say the pending orders will be sent to the 3PL after holding $S(\mathbf{x}_l)$ period.

Case 2. A new order comes in the first $S(\mathbf{x}_l)$ period. Similar to the statement in Case 1, the current pending orders will be held until the period receiving the new order. After receiving the new consecutive order, the system will enter another cluster with different order features $\mathbf{x}_{l+1}$ whose optimal threshold $S(\mathbf{x}_{l+1})$ can be characterized and begin a new cycle. So we have completed the proof of Theorem 1. □

Proof of Proposition 2. By the reformulation of $\bar{q}_{w,M}(\mathbf{x}_l;\Theta,N)$ in (44), we have

$$\bar{q}_{w,M}(\mathbf{x}_l;\Theta,N) = e^{aw}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N), \forall w \in \langle T-1\rangle. \tag{29}$$

According to the Markovian logit model, the probability of making consecutive orders is

$$\begin{aligned}
&\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N) + \bar{q}_{1,M}(\mathbf{x}_l;\Theta,N) + \cdots + \bar{q}_{T-1,M}(\mathbf{x}_l;\Theta,N) \\
&= \bar{q}_{0,M}(\mathbf{x}_l;\Theta,N) + e^{a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N) + \cdots + e^{(T-1)a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N) \\
&= \frac{1-e^{aT}}{1-e^{a}}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)
\end{aligned}$$

where the first equality follows (29) and the second equality is by simple calculation. This completes the proof of Proposition 2. □

Proof of Theorem 2. Before starting the proof, let's first reformulate the transition probabilities $p((\mathbf{x}_l, w), (\mathbf{x}_l, w+1))$ using the choice probabilities of making another order $\{\bar{q}_{\tau,M}(\mathbf{x}_l;\Theta,N)\}_{\tau\in\langle T-1\rangle}$. Recall the original formulation of $p((\mathbf{x}_l, w), (\mathbf{x}_l, w+1)) = \mathbb{P}(w+1;\mathbf{x}_l,\Theta,N)/\mathbb{P}(w;\mathbf{x}_l,\Theta,N)$ in (16) where $\mathbb{P}(w;\mathbf{x}_l,\Theta,N)$ defined in (15) denotes the probability of the event that no new order by time w. Due to the structure of the choice network N, it can be reformulated as $\mathbb{P}(w;\mathbf{x}_l,\Theta,N) = 1-\sum_{\tau=0}^{w-1}\bar{q}_{\tau,M}(\mathbf{x}_l;\Theta,N)$, which leads to the reformulation of transition probabilities as

$$p((\mathbf{x}_l, w), (\mathbf{x}_l, w+1)) = \frac{1-\sum_{\tau=0}^{w}\bar{q}_{\tau,M}(\mathbf{x}_l;\Theta,N)}{1-\sum_{\tau=0}^{w-1}\bar{q}_{\tau,M}(\mathbf{x}_l;\Theta,N)} = 1 - \frac{\bar{q}_{w,M}(\mathbf{x}_l;\Theta,N)}{1-\sum_{\tau=0}^{w-1}\bar{q}_{\tau,M}(\mathbf{x}_l;\Theta,N)}. \tag{30}$$

After the reformulation of transition probabilities, let's define an important function which will play a central role to characterize the close form of personalized thresholds, that is $\delta(w) := h + p((\mathbf{x}_l, w), (\mathbf{x}_l, w+1)) \cdot c - c$. When substituting (30), it becomes

$$\delta(w) = h - \frac{\bar{q}_{w,M}(\mathbf{x}_l;\Theta,N)}{1-\sum_{\tau=0}^{w-1}\bar{q}_{\tau,M}(\mathbf{x}_l;\Theta,N)}c. \tag{31}$$

By the reformulation of $\bar{q}_{w,M}(\mathbf{x}_l;\Theta,N)$ in (29) and (44), we have

$$\begin{aligned}
\bar{q}_{w,M}(\mathbf{x}_l;\Theta,N) &= e^{aw}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N), \forall w \in \langle T-1\rangle \\
\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N) &= \frac{\exp(\boldsymbol{\eta}\mathbf{x}_l + b)}{\sum_{\tau\in\langle T-1\rangle}\exp(\beta_\tau + a\tau) + \sum_{\tau\in\langle T-1\rangle} e^{a\tau}\cdot\exp(\boldsymbol{\eta}\mathbf{x}_l + b)} \\
&= \frac{\exp(\boldsymbol{\eta}\mathbf{x}_l + b)}{\sum_{\tau\in\langle T-1\rangle}\exp(\beta_\tau + a\tau) + \frac{1-e^{aT}}{1-e^{a}}\cdot\exp(\boldsymbol{\eta}\mathbf{x}_l + b)}.
\end{aligned} \tag{32}$$

With the reformulation (29) of $\bar{q}_{w,M}(\mathbf{x}_l;\Theta,N)$ above, we can finally reformulate the function $\delta(w)$ as

$$\delta(w) = h - \frac{e^{aw}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)}{1-\frac{1-e^{aw}}{1-e^{a}}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)}c = h - \frac{1}{\frac{1-\frac{1}{1-e^{a}}\cdot\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)}{e^{aw}\cdot\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)}+\frac{1}{1-e^{a}}}c. \tag{33}$$

With the reformulation of $\delta(w)$, we are ready to characterize close form of the personalized threshold $S(\mathbf{x}_l;\Theta,N)$ case by case (Case (18a), (18b), (19a), (19b) and (19c) in Theorem 2) as follows.

Case (18a) is with conditions that $h < c\cdot(1-e^{a})$ and $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N) < 1-e^{a}$.

For this case we will verify that $S(\mathbf{x}_l;\Theta,N) = (\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil \wedge T)\vee 0$ with the nonlinear function $f_{nl}(\cdot)$ of the close form defined in (17). It is easy to check that $\delta(w)$ is increasing in w for $w\in\mathbb{R}$ by (33) when noting that $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N) < 1-e^{a}$ under **Case (18a)**. Define $f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)) := \min\{w\in\mathbb{R}:\delta(w)\geq 0\}$ as the break point of the sign of $\delta(w)$, we will first verify that $S(\mathbf{x}_l;\Theta,N) = (\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0$ (in **Step 1**) and then characterize the close form of the nonlinear function $f_{nl}(\cdot)$ (in **Step 2**).

Step 1. We verify that $S(\mathbf{x}_l;\Theta,N) = (\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0$ by considering cases as follows depending on $S(\mathbf{x}_l;\Theta,N)=T$ (*Case A*), $S(\mathbf{x}_l;\Theta,N)=0$ (*Case B*) or $1\leq S(\mathbf{x}_l;\Theta,N)\leq T-1$ (*Case C*).

Case A that $S(\mathbf{x}_l;\Theta,N)=T$ which implies that $\delta(T-1) = h+p((\mathbf{x}_l,T-1),(\mathbf{x}_l,T))\cdot c - c = h+p((\mathbf{x}_l,T-1),(\mathbf{x}_l,T))\cdot\Delta_{\mathbf{x}_l}(T)-c<0$ by the characterization of personalized thresholds in (11). By definition, $f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)) > T-1$ which results in $(\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0 = T = S(\mathbf{x}_l;\Theta,N)$. So we have verified that $S(\mathbf{x}_l;\Theta,N) = (\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0$ under *Case A*.

Case B that $S(\mathbf{x}_l;\Theta,N)=0$ which implies that $h+p((\mathbf{x}_l,0),(\mathbf{x}_l,1))\cdot\Delta_{\mathbf{x}_l}(1)-c\geq 0$ by the characterization of personalized thresholds in (11). The recursive definition of $\Delta_{\mathbf{x}_l}(w)$ in (10) implies that

$$\Delta_{\mathbf{x}_l}(w)\leq c, \forall w\in\langle T\rangle \tag{34}$$

which results in $\delta(0) = h+p((\mathbf{x}_l,0),(\mathbf{x}_l,1))\cdot c-c\geq h+p((\mathbf{x}_l,0),(\mathbf{x}_l,1))\cdot\Delta_{\mathbf{x}_l}(1)-c\geq 0$. By definition again, $f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\leq 0$ which results in $(\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0 = 0 = S(\mathbf{x}_l;\Theta,N)$. So we have verified that $S(\mathbf{x}_l;\Theta,N) = (\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0$ under *Case B*.

Case C that $1\leq S(\mathbf{x}_l;\Theta,N)\leq T-1$ which implies that the following two conditions hold by the characterization of personalized thresholds in (11):

$$h+p((\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)),(\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)+1))\cdot\Delta_{\mathbf{x}_l}(S(\mathbf{x}_l;\Theta,N)+1)-c\geq 0, \tag{35}$$

$$h+p((\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)-1),(\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)))\cdot c-c<0, \tag{36}$$

where the second inequality originates from the fact that $\Delta_{\mathbf{x}_l}(S(\mathbf{x}_l;\Theta,N))=c$ by the single dimensional MDP (10) and (35).

To show $S(\mathbf{x}_l;\Theta,N) = (\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0$ under *Case C*, it suffices to show that $S(\mathbf{x}_l;\Theta,N)\leq(\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0$ and $S(\mathbf{x}_l;\Theta,N)\geq(\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0$ hold simultaneously as follows.

First, we will show that $S(\mathbf{x}_l;\Theta,N)\leq(\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0$ which is equivalent to $S(\mathbf{x}_l;\Theta,N)\leq\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\vee 0$ since $S(\mathbf{x}_l;\Theta,N)\leq T$. Prove by contradiction. Assuming that $S(\mathbf{x}_l;\Theta,N)>\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\vee 0$ which is equivalent to $S(\mathbf{x}_l;\Theta,N)>\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil$ due to $S(\mathbf{x}_l;\Theta,N)>0$ under *Case C*. Since $S(\mathbf{x}_l;\Theta,N)$ is an integer, we must have $S(\mathbf{x}_l;\Theta,N)-1\geq\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\geq f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))$. By the definition of $f_{nl}(\cdot)$ and monotonicity of $\delta(w)$ (increasing in w under **Case (18a)**), $\delta(S(\mathbf{x}_l;\Theta,N)-1) = h+p((\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)-1),(\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)))\cdot c-c\geq 0$ which is contradict with (36). So, we have shown that $S(\mathbf{x}_l;\Theta,N)\leq(\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0$.

Second, we will show that $S(\mathbf{x}_l;\Theta,N)\geq(\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T)\vee 0$ which is equivalent to $S(\mathbf{x}_l;\Theta,N)\geq\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil\wedge T$ since $S(\mathbf{x}_l;\Theta,N)\geq 0$. Prove by contradiction again.

Assuming that $S(\mathbf{x}_l;\Theta,N) < \lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil \wedge T$ which is equivalent to $S(\mathbf{x}_l;\Theta,N) < \lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil$ since $S(\mathbf{x}_l;\Theta,N) < T$ under *Case C*. Since $S(\mathbf{x}_l;\Theta,N)$ is an integer, we must have $S(\mathbf{x}_l;\Theta,N) < f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))$. By the definition of $f_{nl}(\cdot)$ and monotonicity of $\delta(w)$ (increasing in w under **Case (18a)**) again, $0 > \delta(S(\mathbf{x}_l;\Theta,N)) = h + p((\mathbf{x}_l, S(\mathbf{x}_l;\Theta,N)),(\mathbf{x}_l, S(\mathbf{x}_l;\Theta,N)+1))\cdot c - c \geq h + p((\mathbf{x}_l, S(\mathbf{x}_l;\Theta,N)),(\mathbf{x}_l, S(\mathbf{x}_l;\Theta,N)+1))\cdot \Delta_{\mathbf{x}_l}(S(\mathbf{x}_l;\Theta,N)+1) - c$ which is contradict with (35), where the last inequality comes from (34). So, we have shown that $S(\mathbf{x}_l;\Theta,N) \geq (\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil \wedge T) \vee 0$. Combine the two steps above together, we have shown that $S(\mathbf{x}_l;\Theta,N) = (\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil \wedge T)\vee 0$ under *Case C*.

Combine *Case A*, *B* and *C* together, we have shown that $S(\mathbf{x}_l;\Theta,N) = (\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil \wedge T)\vee 0$.

Step 2. To characterize the close form of the personalized threshold $S(\mathbf{x}_l;\Theta,N) = (\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil \wedge T)\vee 0$ under **Case (18a)**, it remains to characterize the close form of the nonlinear function $f_{nl}(\cdot)$.

Recall the reformulation of $\delta(w)$ in (33), we have

$$\begin{aligned}
\delta(w)\geq 0 &\Leftrightarrow h\left(1-\frac{1}{1-e^a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)+\frac{1}{1-e^a}\cdot e^{aw}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\right)-c\cdot e^{aw}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\geq 0\\
&\Leftrightarrow h\left(1-\frac{1}{1-e^a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\right)+\left(\frac{h}{1-e^a}-c\right)\cdot e^{aw}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\geq 0\\
&\Leftrightarrow w\geq \frac{1}{a}\left\{\ln(h)-\ln\left(c-\frac{h}{1-e^a}\right)\right\}-\frac{1}{a}\ln(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))+\frac{1}{a}\ln\left(1-\frac{1}{1-e^a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\right)\\
f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N) &= \frac{1}{a}\left\{\ln(h)-\ln\left(c-\frac{h}{1-e^a}\right)\right\}-\frac{1}{a}\ln(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))+\frac{1}{a}\ln\left(1-\frac{1}{1-e^a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\right)
\end{aligned}$$

where the last equivalence comes from the fact that $\frac{h}{1-e^a}-c<0$ under **Case (18a)** and the last equality comes from the definition of $f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)) := \min\{w\in\mathbb{R}:\delta(w)\geq 0\}$. So, we have characterize the close of the nonlinear function $f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)$, so does the personalized threshold $S(\mathbf{x}_l;\Theta,N)$ under **Case (18a)**.

Next, we turn to **Case (18b)**.

Case (18b) is with conditions that $h < c\cdot(1-e^a)$ and $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\geq 1-e^a$.

In this case we wish to verify that $S(\mathbf{x}_l;\Theta,N)=T$. Actually, $\delta(0) = h-\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\cdot c \leq h-(1-e^a)\cdot c < 0$ where the two inequalities originate from conditions in **Case (18b)**. It is easy to check that $\delta(w)$ is decreasing in w for $w\in[0,T-1]$ by the reformulation of $\delta(w)$ in (33) when noting that $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\geq 1-e^a$ under **Case (18b)**. So, $\delta(w)\leq\delta(0)<0$ for $w\in\langle T-1\rangle$, which results in

$$h+p((\mathbf{x}_l,w),(\mathbf{x}_l,w+1))\cdot\Delta_{\mathbf{x}_l}(w+1)-c\leq h+p((\mathbf{x}_l,w),(\mathbf{x}_l,w+1))\cdot c-c=\delta(w)<0,\forall w\in\langle T-1\rangle,$$

where the first inequality comes from the fact that $\Delta_{\mathbf{x}_l}(w)\leq c$ for $w\in\langle T\rangle$ in (34). The conditions above actually imply that $S(\mathbf{x}_l;\Theta,N)=T$ by its characterization in (11). So, we have shown that $S(\mathbf{x}_l;\Theta,N)=T$ under **Case (18b)**.

Now we begin to consider the remaining cases when $h\geq c\cdot(1-e^a)$.

Case (19a) is with conditions that $h\geq c\cdot(1-e^a)$ and $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\leq 1-e^a$.

We wish to verify that $S(\mathbf{x}_l;\Theta,N)=0$ in this case. By the reformulation (33) and the fact that $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\leq 1-e^a$, $\delta(w)$ is increasing in w for $w\in[0,T-1]$ which implies that $\delta(w)\geq\delta(0)\geq 0$ for $w\in[0,T-1]$, where the last inequality comes from the fact that $\delta(0)=h-\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\cdot c\geq h-(1-e^a)\cdot c\geq 0$ under **Case (19a)**. Prove by contradiction to show that $S(\mathbf{x}_l;\Theta,N)=0$ in this case. If $S(\mathbf{x}_l;\Theta,N)>0$, we must have $\delta(S(\mathbf{x}_l;\Theta,N)-1)=h+p((\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)-1),(\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)))\cdot c-c<0$ by (36). Here comes the contradiction since we have shown that $\delta(w)\geq 0$ for $w\in[0,T-1]$. So, we have verified that $S(\mathbf{x}_l;\Theta,N)=0$ under **Case (19a)**.

Case (19b) is with conditions that $h\geq c\cdot(1-e^a)$ and $1-e^a<\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\leq h/c$.

First, we will verify that the personalized threshold $S(\mathbf{x}_l;\Theta,N)$ is either 0 or T in this case, and then characterize the conditions under which it will be one of them. Prove by contradiction to show that $S(\mathbf{x}_l;\Theta,N)$ is either 0 or T. Assuming that $1\le S(\mathbf{x}_l;\Theta,N)\le T-1$ under **Case (19b)**. By the reformulation (33) and the fact that $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)>1-e^a$, $\delta(w)$ is decreasing in w for $w\in[0,T-1]$ under **Case (19b)**. To derive the contradiction, define $w_1:=\lceil\max\{w\in\langle T-1\rangle:\delta(w)\ge 0\}\rceil$ ($w_1\ge 0$ since $\delta(0)=h-\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\cdot c\ge 0$ under **Case (19b)**). We will derive the contradiction of $1\le S(\mathbf{x}_l;\Theta,N)\le T-1$ by considering different cases depending on $w_1=T-1$ (*Case A*) or $w_1<T-1$ (*Case B*) as follows.

Case A that $w_1=T-1$ which implies that $\delta(w)\ge 0$ for $w\in\langle T-2\rangle$ by the definition of w_1 and decreasing of $\delta(w)$. It follows that $\delta(S(\mathbf{x}_l;\Theta,N)-1)=h+p((\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)-1),(\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)))\cdot c-c\ge 0$ since $S(\mathbf{x}_l;\Theta,N)-1\in\langle T-2\rangle$ (since $1\le S(\mathbf{x}_l;\Theta,N)\le T-1$ by assumption), which contradicts (36).

Case B that $w_1<T-1$. We consider two additional cases as follows to derive contradiction of $1\le S(\mathbf{x}_l;\Theta,N)\le T-1$ depending on $S(\mathbf{x}_l;\Theta,N)>w_1$ (*Case B.1*) or $S(\mathbf{x}_l;\Theta,N)\le w_1$ (*Case B.2*).

Case B.1 that $S(\mathbf{x}_l;\Theta,N)>w_1$. By definition of w_1, we must have $0>\delta(S(\mathbf{x}_l;\Theta,N))\ge h+p((\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)),(\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)+1))\cdot\Delta_{\mathbf{x}_l}(S(\mathbf{x}_l;\Theta,N)+1)-c$ which contradicts (35), where the last inequality comes from (34).

Case B.2 that $S(\mathbf{x}_l;\Theta,N)\le w_1$. By the definition of w_1, we must have $0\le\delta(S(\mathbf{x}_l;\Theta,N)-1)=h+p((\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)-1),(\mathbf{x}_l,S(\mathbf{x}_l;\Theta,N)))\cdot c-c$ which contradicts (36).

So, we have verified that the personalized threshold $S(\mathbf{x}_l;\Theta,N)$ is either 0 or T under **Case (19b)**. It remains to characterize the necessary and sufficient conditions under which $S(\mathbf{x}_l;\Theta,N)=T$ which are exactly the conditions stated in (19b) by the characterization of personalized thresholds in (11).

Case (19c) is with conditions that $h\ge c\cdot(1-e^a)$ and $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)>h/c$.

We wish to verify $S(\mathbf{x}_l;\Theta,N)=T$ in this case. By the reformulation (33) and the fact that $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)>1-e^a$, $\delta(w)$ is decreasing in w for $w\in[0,T-1]$ which implies that $\delta(w)\le\delta(0)<0$ for $w\in[0,T-1]$, where the last inequality comes from the fact that $\delta(0)=h-\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\cdot c<0$ under **Case (19c)**. It follows that for any $w\in\langle T-1\rangle$

$$h+p((\mathbf{x}_l,w),(\mathbf{x}_l,w+1))\cdot\Delta_{\mathbf{x}_l}(w+1)-c\le h+p((\mathbf{x}_l,w),(\mathbf{x}_l,w+1))\cdot c-c=\delta(w)<0,$$

which implies that $S(\mathbf{x}_l;\Theta,N)=T$ by its characterization in (11), where the first inequality is due to $\Delta_{\mathbf{x}_l}(w)\le c$ for $w\in\langle T\rangle$ by (34). So, we have verified that $S(\mathbf{x}_l;\Theta,N)=T$ under **Case (19c)**.

That completes the proof of Theorem 2. □

In the remaining part, we will first present Proposition 4 and its proof, based on which Theorem 3 can be proved.

Proposition 4 (Piecewise Linear Approximation of Personalized Thresholds). *Under the Markovian logit model shown in (13) with network N and parameters $\Theta=(\boldsymbol{\eta},\boldsymbol{\beta},a,b)$, denote $X(\epsilon)$ as the region of features $\mathbf{x}_l$ such that their corresponding probability of making another consecutive orders is bounded by ϵ, that is $X(\epsilon)=\{\mathbf{x}_l\in X|\sum_{\tau\in\langle T-1\rangle}\bar{q}_{\tau,M}(\mathbf{x}_l;\Theta,N)\le\epsilon\}$. If in addition $h<c\cdot(1-e^a)$ and $\epsilon<1-e^{aT}$, then for any $\mathbf{x}_l\in X(\epsilon)$ the personalized threshold $S(\mathbf{x}_l;\Theta,N)$ can be bounded by a piecewise-linear function as:*

$$0\le S(\mathbf{x}_l;\Theta,N)-(\lceil f_l(\mathbf{x}_l;\Theta)\rceil\wedge T)\vee 0\le\left\lfloor 1+\frac{1}{a}\ln\left(1+\frac{\epsilon\wedge\epsilon_U}{1-\epsilon\wedge\epsilon_U}\right)+\frac{1}{a}\ln\left(1-\frac{\epsilon\wedge\epsilon_U}{1-e^{aT}}\right)\right\rfloor,\qquad(37)$$

where the linear function $f_l(\cdot)$ and scalar ϵ_U are defined in (20) and (21).

Proof of Proposition 4. Firstly, we will verify that the personalized threshold $S(\mathbf{x}_l;\Theta,N)$ shares the close form in (18a) with conditions of Proposition 4. For any $\mathbf{x}_l\in X(\epsilon)$, we have $\sum_{\tau\in\langle T-1\rangle}\bar{q}_{\tau,M}(\mathbf{x}_l;\Theta,N)=\frac{1-e^{aT}}{1-e^a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\le\epsilon<1-e^{aT}$ which implies that $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)<1-e^a$, where the equality is obtained by substituting (29). Together with the assumption that $h<c\cdot$

$(1-e^a)$, the personalized threshold share the close form $S(\mathbf{x}_l;\Theta,N) = (\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil \wedge T)\vee 0$ for any $\mathbf{x}_l \in X(\epsilon)$ by Theorem 2. So, to approximate the personalized threshold $S(\mathbf{x}_l;\Theta,N)$, it suffices to approximate the nonlinear function $f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))$. We will firstly bound the nonlinear function $f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))$ with a linear counterpart $f_l(\mathbf{x}_l;\Theta)$ (in **Step 1**) and then show that this will result in the piecewise-linear approximation (37) of the personalized threshold (in **Step 2**).

Step 1. Bound $f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))$ with $f_l(\mathbf{x}_l;\Theta)$.

Recall the reformulation of $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)$ in (32) and the definition of $f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))$ in (17), we can reformulate the nonlinear function as

$$
\begin{aligned}
f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)) =& \frac{1}{a}\left\{\ln(h) - \ln\left(c - \frac{h}{1-e^a}\right)\right\} + \frac{1}{a}\cdot\ln\left(1 - \frac{1}{1-e^a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\right) \\
& - \frac{1}{a}\cdot\left\{\boldsymbol{\eta}\mathbf{x}_l + b - \ln\left(\sum_{\tau=0}^{T-1}\exp(\beta_\tau + a\tau) + \frac{1-e^{aT}}{1-e^a}\cdot\exp(\boldsymbol{\eta}'\mathbf{x}_l + b)\right)\right\} \\
=& \frac{1}{a}\left\{\ln(h) - \ln\left(c - \frac{h}{1-e^a}\right)\right\} + \frac{1}{a}\cdot\ln\left(1 - \frac{1}{1-e^a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\right) \\
& - \frac{1}{a}\cdot\left\{\boldsymbol{\eta}\mathbf{x}_l + b - \ln\left(\sum_{\tau=0}^{T-1}\exp(\beta_\tau + a\tau)\right) - \ln\left(1 + \frac{\frac{1-e^{aT}}{1-e^a}\cdot\exp(\boldsymbol{\eta}\mathbf{x}_l + b)}{\sum_{\tau=0}^{T-1}\exp(\beta_\tau + a\tau)}\right)\right\} \\
=& f_l(\mathbf{x}_l;\Theta) + \frac{1}{a}\ln\left(1 + \frac{\frac{1-e^{aT}}{1-e^a}\cdot\exp(\boldsymbol{\eta}\mathbf{x}_l + b)}{\sum_{\tau=0}^{T-1}\exp(\beta_\tau + a\tau)}\right) + \frac{1}{a}\cdot\ln\left(1 - \frac{1}{1-e^a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\right) \\
=& f_l(\mathbf{x}_l;\Theta) + \frac{1}{a}\ln\left(1 + \frac{\epsilon'(\mathbf{x}_l;\Theta)}{1-\epsilon'(\mathbf{x}_l;\Theta)}\right) + \frac{1}{a}\ln\left(1 - \frac{1}{1-e^{aT}}\epsilon'(\mathbf{x}_l;\Theta)\right)
\end{aligned}
$$

where the first equality is obtained by substituting the reformulation of $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)$ in (32) into the definition of $f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))$ in (17), the second and third equality are by simple calculation, and the finally equality originates from the the definition of $\epsilon'(\mathbf{x}_l;\Theta) := \frac{1-e^{aT}}{1-e^a}\cdot\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N) \le \epsilon$.

Define the function $g(\epsilon') := \frac{1}{a}\ln\left(1 + \frac{\epsilon'}{1-\epsilon'}\right) + \frac{1}{a}\ln\left(1 - \frac{1}{1-e^{aT}}\epsilon'\right)$ for $\epsilon' \in [0,\epsilon]$, then $f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)) = f_l(\mathbf{x}_l;\Theta) + g(\epsilon'(\mathbf{x}_l;\Theta))$ by the reformulation above. The function can be reformulated as $g(\epsilon') = \frac{1}{a}\ln\left(\frac{1}{1-e^{aT}}\cdot\left(1 - \frac{e^{aT}}{1-\epsilon'}\right)\right)$ which implies that it is increasing in ϵ'. This will result in the bound $0 = g(0) \le g(\epsilon'(\mathbf{x}_l;\Theta)) \le g(\epsilon)$ which implies the bound for the nonlinear function $f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))$ for any $\mathbf{x}_l \in X(\epsilon)$:

$$
0 \le f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)) - f_l(\mathbf{x}_l;\Theta) \le g(\epsilon) = \frac{1}{a}\ln\left(1 + \frac{\epsilon}{1-\epsilon}\right) + \frac{1}{a}\ln\left(1 - \frac{1}{1-e^{aT}}\epsilon\right). \tag{38}
$$

With the bound in (38) between $f_{nl}(\cdot)$ and $f_l(\cdot)$, we are ready to characterize the piecewise-linear approximation (37) of the personalized threshold in **Step 2**.

Step 2. Piecewise linear approximation for the personalized threshold $S(\mathbf{x}_l;\Theta,N)$.

Divide the set $X(\epsilon)$ into two disjoint subsets: $X(\epsilon) = X_1(\epsilon) \cup X_2(\epsilon)$, where

$$
X_1(\epsilon) := \left\{x \in X(\epsilon) \,\middle|\, \sum_{\tau=0}^{T-1}\bar{q}_{\tau,M}(\mathbf{x}_l;\Theta,N) > \epsilon_U\right\},
$$
$$
X_2(\epsilon) := \left\{x \in X(\epsilon) \,\middle|\, \sum_{\tau=0}^{T-1}\bar{q}_{\tau,M}(\mathbf{x}_l;\Theta,N) \le \epsilon_U\right\}.
$$

To show the bound (37) holds for any $\mathbf{x}_l \in X(\epsilon)$, it suffices to show the bound holds for any $\mathbf{x}_l \in X_1(\epsilon)$ and $\mathbf{x}_l \in X_2(\epsilon)$ respectively.

For any $\mathbf{x}_l \in X_2(\epsilon)$, we must have

$$
\begin{aligned}
0 \le& S(\mathbf{x}_l;\Theta,N) - (\lceil f_l(\mathbf{x}_l;\Theta)\rceil \wedge T)\vee 0 \\
=& (\lceil f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N))\rceil \wedge T)\vee 0 - (\lceil f_l(\mathbf{x}_l;\Theta)\rceil \wedge T)\vee 0
\end{aligned}
$$

$$\leq \lfloor 1 + f_{nl}(\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)) - f_l(\mathbf{x}_l;\Theta) \rfloor$$
$$\leq \left\lfloor 1 + \frac{1}{a}\ln\left(1 + \frac{\epsilon\wedge\epsilon_U}{1-\epsilon\wedge\epsilon_U}\right) + \frac{1}{a}\ln\left(1 - \frac{\epsilon\wedge\epsilon_U}{1-e^{aT}}\right)\right\rfloor$$

which is exactly the bound (37), where the second inequality comes from Lemma 3 and the last inequality originates form the bound (38) when noting that $X_2(\epsilon)\subseteq X(\epsilon\wedge\epsilon_U)$.

It remains to verify the bound for any $\mathbf{x}_l\in X_1(\epsilon)$. Denote the deterministic utility of purchasing another order as $u(\mathbf{x}_l)=\boldsymbol{\eta}'\mathbf{x}_l+b$, then both $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)$ and $f_l(\mathbf{x}_l;\Theta)$ can be written as functions of $u(\mathbf{x}_l)$ by their definitions in (32) and (20) as follows:

$$\bar{q}'_{0,M}(u(\mathbf{x}_l);\Theta,N) = \frac{\exp(u(\mathbf{x}_l))}{\sum_{\tau=0}^{T-1}\exp(\beta_\tau + a\tau) + \frac{1-e^{aT}}{1-e^a}\exp(u(\mathbf{x}_l))}, \tag{39}$$

$$f'_l(u(\mathbf{x}_l);\Theta) := -\frac{1}{a}u(\mathbf{x}_l) + C_0, \tag{40}$$

both of which are strictly increasing in $u(\mathbf{x}_l)$. Then, for any $\mathbf{x}_l\in X_1(\epsilon)$, we must have $\frac{1-e^{aT}}{1-e^a}\cdot\bar{q}'_{0,M}(u(\mathbf{x}_l);\Theta,N)>\epsilon_U=\frac{1-e^{aT}}{1-e^a}\cdot\bar{q}'_{0,M}(-a(T-1-C_0);\Theta,N)$ (the "=" here is by definition of ϵ_U) which implies that $u(\mathbf{x}_l)>-a(T-1-C_0)$ by the increasing of $\bar{q}'_{0,M}(u;\Theta,N)$ in u. This will lead to $f'_l(u(\mathbf{x}_l);\Theta)>f'_l(-a(T-1-C_0);\Theta)=T-1$ since $f'_l(u;\Theta)$ is also strictly increasing in u. It follows that $0\leq S(\mathbf{x}_l;\Theta,N)-(\lceil f_l(\mathbf{x}_l;\Theta)\rceil\wedge T)\vee 0 = S(\mathbf{x}_l;\Theta,N)-(\lceil f'_l(u(\mathbf{x}_l);\Theta)\rceil\wedge T)\vee 0 = S(\mathbf{x}_l;\Theta,N)-T$ (the "$\leq$" comes from (38)), which implies that $S(\mathbf{x}_l;\Theta,N)=T$ and $S(\mathbf{x}_l;\Theta,N)-(\lceil f_l(\mathbf{x}_l;\Theta)\rceil\wedge T)\vee 0=0$. So we have verified that the bound (37) also holds for any $\mathbf{x}_l\in X_1(\epsilon)$, which completes the proof of Proposition 4. □

Proof of Theorem 3. By Proposition 4, the bound in Theorem 3 holds for any features $\mathbf{x}_l\in X(\epsilon)$ with any $\epsilon<1-e^{aT}$. It remains to show it still holds for any $\mathbf{x}_l\in X^c$, where the set X^c is defined as

$$X^c = \left\{\mathbf{x}_l\in X \,\middle|\, \sum_{\tau\in\langle T-1\rangle}\bar{q}_{\tau,M}(\mathbf{x}_l;\Theta,N)\geq 1-e^{aT}\right\}$$
$$= \left\{\mathbf{x}_l\in X \,\middle|\, \frac{1-e^{aT}}{1-e^a}\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\geq 1-e^{aT}\right\}$$

where the second equality is obtained by substituting (29). Then, we have $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)\geq 1-e^a$ for any $\mathbf{x}_l\in X^c$ by definition and $h<c\cdot(1-e^a)$ by the assumption that $h<c\cdot(1-e^a)/(1+e^a)$, which implies that $S(\mathbf{x}_l;\Theta,N)=T$ by (18b) in Theorem 2. To verify the bound in Theorem 3, it suffices to show that the piecewise-linear function $f_{pl}(\mathbf{x}_l)=(\lceil f_l(\mathbf{x}_l;\Theta)\rceil\wedge T)\vee 0$ equals T for any $\mathbf{x}_l\in X^c$ as follows.

Recall that both $\bar{q}_{0,M}(\mathbf{x}_l;\Theta,N)$ and $f_l(\mathbf{x}_l;\Theta)$ can be equivalently written as $\bar{q}'_{0,M}(\mathbf{x}_l;\Theta,N)$ (in (39)) and $f'_l(\mathbf{x}_l;\Theta)$ (in (40)) respectively, which are strictly increasing functions of $u(\mathbf{x}_l)=\boldsymbol{\eta}'x+b$. By definition of the set $\mathbf{x}_l\in X^c$, it holds that $\frac{1-e^{aT}}{1-e^a}\bar{q}'_{0,M}(u(\mathbf{x}_l);\Theta,N)\geq 1-e^{aT}$. By simple calculation, it is easy to verify that $1-e^{aT}>\epsilon_U$ under the imposed assumption that $h<c\cdot(1-e^a)/(1+e^a)$, which implies that $\frac{1-e^{aT}}{1-e^a}\bar{q}'_{0,M}(u(\mathbf{x}_l);\Theta,N)>\epsilon_U=\frac{1-e^{aT}}{1-e^a}\cdot\bar{q}'_{0,M}(-a(T-1-C_0);\Theta,N)$ (the "=" here is by definition of ϵ_U). This will result in $u(\mathbf{x}_l)>-a(T-1-C_0)$ by increasing of the function $\bar{q}'_{0,M}(u;\Theta,N)$ in u. It follows that $f'_l(u(\mathbf{x}_l);\Theta)>f'_l(-a(T-1-C_0);\Theta)=T-1$ by increasing of the function $f'_l(u;\Theta)$ in u, which implies that $f_{pl}(\mathbf{x}_l)=(\lceil f_l(\mathbf{x}_l;\Theta)\rceil\wedge T)\vee 0=(\lceil f'_l(u(\mathbf{x}_l);\Theta)\rceil\wedge T)\vee 0=T=S(\mathbf{x}_l;\Theta,N)$. So we have completed the proof of Theorem 3. □

LEMMA 3. *For any scalars A and B with $A\geq B$, it holds that* $0\leq(\lceil A\rceil\wedge T)\vee 0-(\lceil B\rceil\wedge T)\vee 0\leq\lfloor A-B+1\rfloor$.

Proof of Lemma 3. To prove Lemma 3, it suffices to prove the bound that $(\lceil A\rceil\wedge T)\vee 0-(\lceil B\rceil\wedge T)\vee 0\leq A-B+1$ when noting that both $(\lceil A\rceil\wedge T)\vee 0$ and $(\lceil B\rceil\wedge T)\vee 0$ are integers. The bound will be verified by considering different cases depending on $A \le 0$ (**Case 1**), $A \ge T$ (**Case B**), or $0 < A < T$ (**Case C**).

Case A that $A \le 0$ which implies that $B \le A \le 0$. This will lead to $(\lceil A \rceil \wedge T) \vee 0 - (\lceil B \rceil \wedge T) \vee 0 = 0 - 0 \le A - B + 1$.

Case B that $A \ge T$ which implies that

$$(\lceil A \rceil \wedge T) \vee 0 - (\lceil B \rceil \wedge T) \vee 0 = T - (\lceil B \rceil \wedge T) \vee 0 = \begin{cases} 0 \le A - B + 1, \text{ if } B > T - 1, \\ T - \lceil B \rceil \vee 0 \le A - B + 1, \textbf{ if } B \le T - 1. \end{cases} \tag{41}$$

Case C that $0 < A < T$ which implies that $B \le A < T$. This will lead to

$$(\lceil A \rceil \wedge T) \vee 0 - (\lceil B \rceil \wedge T) \vee 0 = \lceil A \rceil - \lceil B \rceil \vee 0 = \begin{cases} \lceil A \rceil - \lceil B \rceil \le A - B + 1, \text{ if } B > 0, \\ \lceil A \rceil \le A + 1 \le A - B + 1, \text{ if } B \le 0. \end{cases} \tag{42}$$

This completes the proof of Lemma 3. □

Proof of Proposition 3. To prove Proposition 3, it suffices to show that each term $\ln(\Pr(\mathbf{x}_l^k, DT_l^k; \Theta))$ in (23) is concave in parameters $\Theta = (\boldsymbol{\eta}, \boldsymbol{\beta}, a, b)$ for any $k \in [K]$ and $l \in [L_k]$. By (13), we have

$$\begin{aligned} \Pr(\mathbf{x}_l^k, DT_l^k; \Theta) &= \prod_{w=0}^{DT_l^k - 2} \bar{p}_{w,w+1}(\mathbf{x}_l^k; \Theta, N) \cdot \bar{p}_{DT_l^k - 1, y_l^k}(\mathbf{x}_l^k; \Theta, N) \\ &= \prod_{w=0}^{DT_l^k - 2} \frac{\exp\left(a + g_{w+1}(\mathbf{x}_l^k; \Theta, N)\right)}{\exp(g_w(\mathbf{x}^k; \Theta, N))} \cdot \frac{\mathcal{I}(y_l^k = \triangle) \cdot \exp(\beta_{DT_l^k - 1}) + \mathcal{I}(y_l^k = M) \cdot \exp(\boldsymbol{\eta}' \mathbf{x}_l^k + b)}{\exp(g_{DT_l^k - 1}(\mathbf{x}_l^k; \Theta, N))} \\ &= \frac{\mathcal{I}(y_l^k = \triangle) \cdot \exp(\beta_{DT_l^k - 1} + a(DT_l^k - 1)) + \mathcal{I}(y_l^k = M) \cdot \exp(\boldsymbol{\eta}' \mathbf{x}_l^k + b + a(DT_l^k - 1))}{\exp(g_0(\mathbf{x}_l^k; \Theta, N))}. \end{aligned} \tag{43}$$

It follows that $\ln(\Pr(\mathbf{x}_l^k, DT_l^k; \Theta)) = \mathcal{I}(y_l^k = \triangle) \cdot (\beta_{DT_l^k - 1} + a(DT_l^k - 1)) + \mathcal{I}(y_l^k = M) \cdot (\boldsymbol{\eta}' \mathbf{x}_l^k + b + a(DT_l^k - 1)) - g_0(\mathbf{x}_l^k; \Theta, N)$, where the first term is linear in Θ. Hence, it suffices to show that $g_0(\mathbf{x}_l^k; \Theta, N)$ is convex in Θ. By Lemma 5, $g_0(\mathbf{x}_l^k; \Theta, N)$ can be reformulated as $g_0(\mathbf{x}; \Theta, N) = \ln\left(\sum_{\tau=0}^{T-1} \exp\left(\beta_\tau + a\tau\right) + \exp(\boldsymbol{\eta}\mathbf{x} + b + a\tau)\right)$, whose convexity in Θ is confirmed by the following Lemma 4. This completes the proof. □

LEMMA 4 (CONVEXITY OF LOG EXPONENTIAL FUNCTIONS). *The function* $g(\mathbf{x}) = \ln(\sum_{i=1}^n p_i \exp(\mathbf{b}_i' \mathbf{x} + c_i))$ *with* $p_i \ge 0$ *for* $i = 1, 2, \dots, n$ *is a convex function in* $\mathbf{x}$.

Proof of Lemma 4. The convexity of function $g(\mathbf{x})$ in $\mathbf{x}$ is equivalent to the fact that the epigraph, $\text{epi}g(\mathbf{x}) = \{(\mathbf{x}, y) | g(\mathbf{x}) \le y\}$, is a convex set. Noting that $\text{epi}g(\mathbf{x})$ can be equivalent reformulated as $\text{epi}g(\mathbf{x}) = \{(\mathbf{x}, y) | \sum_{i=1}^n p_i \exp(\mathbf{b}_i' \mathbf{x} + c_i - y) \le 0\}$ which is a convex set since $f(\mathbf{x}, y) = \sum_{i=1}^n p_i \exp(\mathbf{b}_i' \mathbf{x} + c_i - y)$ is jointly convex in $(\mathbf{x}, y)$. This completes the proof of Lemma 4. □

LEMMA 5 (REFORMULATE $g_0(\mathbf{x}; \Theta)$). *The function* $g_w(\mathbf{x}; \Theta, N)$ *defined in (13) can be reformulated as*

$$g_w(\mathbf{x}; \Theta, N) = \ln\left(\sum_{\tau=w}^{T-1} \exp\left(\beta_\tau + a(\tau - w)\right) + \exp\left(\boldsymbol{\eta}\mathbf{x} + b + a(\tau - w)\right)\right), \forall w \in \langle T - 1 \rangle.$$

Proof of Lemma 5. Prove by induction on $w \in \langle T - 1 \rangle$. By (13), we have $g_{T-1}(\mathbf{x}_l; \Theta, N) = \ln(\exp(\beta_{T-1}) + \exp\left(\boldsymbol{\eta}\mathbf{x}_l + b\right))$ which is exactly the form of $g_{T-1}(\mathbf{x}_l; \Theta, N)$ we want to prove. Now assume that $g_{w+1}(\mathbf{x}_l; \Theta, N)$ shares the form in Lemma 5, it remains to show that so does $g_w(\mathbf{x}_l; \Theta, N)$. By (13) again, we have

$$\begin{aligned} g_w(\mathbf{x}_l; \Theta, N) &= \ln\left(\exp(\beta_w) + \exp(\boldsymbol{\eta}\mathbf{x}_l + b) + \exp\left(a + g_{w+1}(\mathbf{x}_l; \Theta, N)\right)\right) \\ &= \ln\left(\exp(\beta_w) + \exp(\boldsymbol{\eta}\mathbf{x}_l + b) + \exp(a) \cdot \left(\sum_{\tau=w+1}^{T-1} \exp\left(\beta_\tau + a(\tau - w - 1)\right) + \exp(\boldsymbol{\eta}\mathbf{x}_l + b + a(\tau - w - 1))\right)\right) \end{aligned}$$

$$= \ln\left(\sum_{\tau=w}^{T-1} \exp\left(\beta_\tau + a(\tau - w)\right) + \exp(\boldsymbol{\eta}\mathbf{x}_l + b + a(\tau - w))\right),$$

where the second equality is obtained by substituting the induction assumption of $g_{w+1}(\mathbf{x}_l;\Theta,N)$ into the first equality. So we have completed the induction proof of Lemma 5. □

LEMMA 6 (EQUIVALENCE OF MLE AND THE MNL CHOICE MODEL). *There exists an MNL model with choice alternatives in the set $\mathcal{L}$ and the deterministic utility terms being linear in the order features $\mathbf{x}$, where $\mathcal{L} = \{(0,\triangle),(0,M),(1,\triangle),(1,M),\cdots,(T-1,\triangle),(T-1,M)\}$ is defined as a set that includes all the arcs to the terminal node Δ or M in network N. Under this MNL model, the choice probability of each alternative is the same as that of the corresponding arc in network N under MLM.*

Proof of Lemma 6. Recall that the MLM model with network N and parameters $\Theta = (\boldsymbol{\eta},\boldsymbol{\beta},a,b)$ assumes that the deterministic utility term is a linear function of features $\mathbf{x}_l$, and random error terms follow i.i.d. Gumbel distributions. Now we are going to construct a MNL model with alternatives in $\mathcal{L}$ such that it shares the same unconditional choice probability of each arc in $\mathcal{L}$ as the MLM model.

We construct the MNL model by further reformulating the unconditional choice probability $\bar{q}_{w,y}(\mathbf{x}_l;\Theta,N)$ of arc $(w,y)\in\mathcal{L}$ in the MLM model as follows. Firstly, by the close from of choice probabilities for the MLM model in (13), we have

$$\begin{aligned}
\bar{q}_{w,y}(\mathbf{x}_l;\Theta,N) &= \prod_{\tau=0}^{w-1} \bar{p}_{\tau,\tau+1}(\mathbf{x}_l;\Theta,N)\cdot \bar{p}_{w,y}(\mathbf{x}_l;\Theta,N)\\
&= \prod_{\tau=0}^{w-1} \frac{\exp\left(a + g_{\tau+1}(\mathbf{x}_l;\Theta,N)\right)}{\exp(g_\tau(\mathbf{x}_l;\Theta,N))}\cdot \frac{\mathcal{I}(y=\triangle)\cdot\exp(\beta_w) + \mathcal{I}(y=M)\cdot\exp(\boldsymbol{\eta}\mathbf{x}_l + b)}{\exp(g_w(\mathbf{x}_l;\Theta,N))}\\
&= \frac{\mathcal{I}(y=\triangle)\cdot\exp(\beta_w + aw) + \mathcal{I}(y=M)\cdot\exp(\boldsymbol{\eta}\mathbf{x}_l + b + aw)}{\exp(g_0(\mathbf{x}_l;\Theta,N))}.
\end{aligned}$$

By substituting the reformulation of $\exp(g_0(\mathbf{x}_l;\Theta,N))$ in Lemma 5, above expression of $\bar{q}_{w,y}(\mathbf{x}_l;\Theta,N)$ turns to

$$\bar{q}_{w,y}(\mathbf{x}_l;\Theta,N) = \frac{\mathcal{I}(y=\triangle)\cdot\exp(\beta_w + aw) + \mathcal{I}(y=M)\cdot\exp(\boldsymbol{\eta}\mathbf{x}_l + b + aw)}{\sum_{\tau=0}^{T-1}\exp\left(\beta_\tau + a\tau\right) + \exp(\boldsymbol{\eta}\mathbf{x}_l + b + a\tau)}, \tag{44}$$

which is exactly the expression of choice probabilities for the MNL model that we are familiar with. The MNL model we are trying to construct is clear by setting the deterministic utility term of the alternative (w,y) in labels set $\mathcal{L}$ as the linear function $\mathcal{I}(y=\triangle)\cdot(\beta_w + aw) + \mathcal{I}(y = M)\cdot(\boldsymbol{\eta}\mathbf{x}_l + b + aw)$. The construction here confirms that the MLM and MNL model share the same choice probabilities of arcs in the label set $\mathcal{L}$, which completes the proof of Lemma 6. □

LEMMA 7. *There exists a logistic regression model that is equivalent to the MLE method of MLM.*

Proof of Lemma 7. Note that the logistic regression model is an MLE method of a *multinomial logit* (MNL) choice model. Hence, it is sufficient to show the equivalence between MLM and the MNL choice model. Hence Lemma 7 is a straightforward result from Lemma 6. □

Proof of Proposition 5. Denote the denominator of (45) as $\hat{\mathbb{P}}(\mathbf{x},w)$, the estimator in (45) can be rewritten as

$$\hat{\mathbb{P}}(\tau|w;\mathbf{x}) = \sum_{\mathbf{x}',w',\tau':(w',\tau')\in\mathcal{T}_{\mathbf{x}'},\mathbf{x}'\in\mathcal{X}} \frac{\mathcal{K}_1(w,w';\zeta_1)\cdot \prod_{i=3,\dots,d_x+2}\mathcal{K}_i(x_i,x_i';\zeta_i)}{\hat{\mathbb{P}}(\mathbf{x},w)}\cdot\mathcal{K}_2(\tau,\tau';\zeta_2),$$

where the kernel function for the categorical variable is defined as

$$\mathcal{K}_2(\tau,\tau';\zeta_2) = \begin{cases} 1-\zeta_2, & \text{if } \tau = \tau',\\ \zeta_2/T, & \text{if } \tau\neq\tau'.\end{cases}$$

With the above reformulation of the estimator $\hat{\mathbb{P}}(\tau|w;\mathbf{x})$, we have

$$\begin{aligned}\sum_{\tau=0}^{T}\hat{\mathbb{P}}(\tau|w;\mathbf{x}) &= \sum_{\mathbf{x}',w',\tau':(w',\tau')\in\mathcal{T}_{\mathbf{x}'},\mathbf{x}'\in\mathcal{X}} \frac{\mathcal{K}_1(w,w';\zeta_1)\cdot\prod_{i=3,\ldots,d_x+2}\mathcal{K}_i(x_i,x_i';\zeta_i)}{\hat{\mathbb{P}}(\mathbf{x},w)}\cdot\left\{\sum_{\tau=0}^{T}\mathcal{K}_2(\tau,\tau';\zeta_2)\right\}.\\ &= \sum_{\mathbf{x}',w',\tau':(w',\tau')\in\mathcal{T}_{\mathbf{x}'},\mathbf{x}'\in\mathcal{X}} \frac{\mathcal{K}_1(w,w';\zeta_1)\cdot\prod_{i=3,\ldots,d_x+2}\mathcal{K}_i(x_i,x_i';\zeta_i)}{\hat{\mathbb{P}}(\mathbf{x},w)}\cdot\left\{1-\zeta_2+\frac{\zeta_2}{T}\cdot T\right\}.\\ &= 1,\end{aligned}$$

where the second equality follows the definition of the kernel function $\mathcal{K}_2(\cdot,\cdot;\zeta_2)$ and the last equality is by the definition of $\hat{\mathbb{P}}(\mathbf{x},w)$. The non-negativity of $\hat{\mathbb{P}}(\tau|w;\mathbf{x})$ follows from the non-negativity of kernel functions which completes the proof of Proposition 5. ☐

B. The KCDE and RF Methods

B.1. Using KCDE to Estimate The Transition Probabilities and The Probability of Placing Consecutive Orders

According to Section 6, we obtain a set of transaction paths from the data. However, to train the KCDE method we need a sequence of observed transitions. This requires us to first translate the transaction paths to the transitions among states with no new orders. Note that each entry in $path^k=\{(\mathbf{x}_l^k,DT_l^k)\}_{l\in[L_k]}$ corresponds to a sequence of transitions in a set of single-dimensional MDP problems. Specifically, each entry $(\mathbf{x}_l^k,DT_l^k)$, $l\in[L_k-1]$, means that upon receiving an order featured by $\mathbf{x}_l^k$, a new order arrives after DT_l^k periods. This corresponds to a set of transitions: $\{((\mathbf{x}_l^k,w),(\mathbf{x}_l^k,w+1))\}_{w\in\langle DT_l^k-2\rangle}\cup\{((\mathbf{x}_l^k,DT_l^k-1),(\mathbf{x}_{l+1}^k,0))\}$. For $l=L_k$, the consumer has made the maximum number of consecutive orders and she will not place any order within T periods. This corresponds to the transitions: $\{((\mathbf{x}_l^k,w),(\mathbf{x}_l^k,w+1))\}_{w\in\langle T-1\rangle}$. Let $\mathcal{T}_{\mathbf{x}}$ denote a set of observed transitions in the single-dimensional MDP problem defined in the cluster identified by $\mathbf{x}$. Let $\mathcal{X}$ denote a set of observed order features. The sets $\mathcal{T}_{\mathbf{x}}$, $\mathbf{x}\in\mathcal{X}$, serve as standard training data for the KCDE method.

Our goal is to estimate the transition probabilities of the MDP problem defined by (5) to (9). According to Theorem 1, it suffices to estimate $p((\mathbf{x}_l,w),(\mathbf{x}_l,w+1))$, $w\in\langle T-1\rangle$, for each single-dimensional MDP. By definition, the conditional probability $p((\mathbf{x}_l,w),(\mathbf{x}_l,w+1))$ is equal to the probability that no new order arrives for $w+1$ periods divided by the probability that no new order arrives for w periods after receiving an order featured by $\mathbf{x}_l$.

We represent the probability of each event using kernel functions. Note that for each order featured by $\mathbf{x}_l$, the number of periods without new order arrivals is a categorical variable. According to the KCDE method (Hall et al. 2004), for any two values w and τ, the probability that the categorical variable equals w and τ before and after a transition, respectively, in a single-dimensional MDP featured by $\mathbf{x}_l$ is $\frac{1}{|\mathcal{T}_{\mathbf{x}_l}|}\sum_{w',\tau':(w',\tau')\in\mathcal{T}_{\mathbf{x}_l}}\mathcal{K}_1(w,w';\zeta_1)\mathcal{K}_2(\tau,\tau';\zeta_2)$, where ζ_i represents the bandwidth for $i=1,2$. Here $|\mathcal{T}_{\mathbf{x}_l}|$ denotes the cardinality of $\mathcal{T}_{\mathbf{x}_l}$, which equals the number of joint observations of (w',τ') before and after a transition in a single-dimensional MDP featured by $\mathbf{x}_l$.

The kernel function for categorical variables with a total of $T+1$ categories is defined as $\mathcal{K}_i(w,w';\zeta_i)=(\zeta_i/T)^{\mathcal{I}(w'\neq w)}(1-\zeta_i)^{1-\mathcal{I}(w'\neq w)}$, $i=1,2$ (see Equation (4) in Hall et al. (2004)). Bowman (1980) shows that this kernel type of probability density function with bandwidths estimated from an MLE method converges (in probability) to the true density function. By definition, the conditional probability for a single-dimensional MDP to transit from w to τ can be calculated as $\sum_{w',\tau':(w',\tau')\in\mathcal{T}_{\mathbf{x}_l}}\mathcal{K}_1(w,w';\zeta_1)\cdot\mathcal{K}_2(\tau,\tau';\zeta_2)\ /\sum_{w':(w',\tau')\in\mathcal{T}_{\mathbf{x}_l}}\mathcal{K}_1(w,w';\zeta_1)$.

The data includes various single-dimensional MDPs. Thus, the observed (w',τ') may be associated with different single-dimensional MDPs. We weight each observation by the kernel density function of its associated single-dimensional MDP featured by $\mathbf{x}$. In particular, we define the conditional probability of a single-dimensional MDP featured by $\mathbf{x}$ to transit from w to τ as follows:

$$\hat{\mathbb{P}}(\tau|w;\mathbf{x}) = \frac{\sum_{\mathbf{x}',w',\tau':(w',\tau')\in\mathcal{T}_{\mathbf{x}'},\mathbf{x}'\in\mathcal{X}} \mathcal{K}_1(w,w';\zeta_1)\cdot\mathcal{K}_2(\tau,\tau';\zeta_2)\cdot\prod_{i=3,\ldots,d_x+2}\mathcal{K}_i(x_i,x_i';\zeta_i)}{\sum_{\mathbf{x}',w':(w',\tau')\in\mathcal{T}_{\mathbf{x}'},\mathbf{x}'\in\mathcal{X}} \mathcal{K}_1(w,w';\zeta_1)\cdot\prod_{i=3,\ldots,d_x+2}\mathcal{K}_i(x_i,x_i';\zeta_i)}. \tag{45}$$

Note that the weight defined by the product of kernel functions $\prod_{i=3,\ldots,d_x+2}\mathcal{K}_i(x_i,x_i';\zeta_i)$ represents the similarity of the two single-dimensional MDPs featured by $\mathbf{x}$ and $\mathbf{x}'$. We assume that the more similar the two order features, the more similar the evolutions of the corresponding single-dimensional MDPs.

We follow the convention to define a kernel function $\mathcal{K}_i(x_i,x_i';\zeta_i)$ for numerical variables as $\mathcal{K}_i(x_i,x_i';\zeta_i)=\mathbb{K}((x_i-x_i')/\zeta_i)$, where $\mathbb{K}(\cdot):\mathbb{R}\to\mathbb{R}_+$ satisfies $\int\mathbb{K}(s)ds<\infty$. The following proposition ensures that the proposed conditional probability in (45) is well defined.

Proposition 5. *The conditional probability defined in (45) satisfies* $\sum_{\tau=0,\ldots,T}\hat{\mathbb{P}}(\tau|w;\mathbf{x})=1$ *and* $\hat{\mathbb{P}}(\tau|w;\mathbf{x})\geq 0$ *for* $\tau=0,\ldots,T$.

We apply the likelihood cross-validation method (Li and Racine 2007) to select bandwidths for mixed data types (containing both categorical and numerical features). Specifically, the bandwidths are selected to maximize the log-likelihood function $\sum_{\mathbf{x},w,\tau:(w,\tau)\in\mathcal{T}_{\mathbf{x}},\mathbf{x}\in\mathcal{X}} z_{\mathbf{x},w,\tau}\log\left(\hat{\mathbb{P}}_{-1}(\tau|w;\mathbf{x})\right)$, where $z_{\mathbf{x},w,\tau}$ is an indicator variable that equals 1 if we observe a transition from w to τ with the order features $\mathbf{x}$ in the data, and 0 otherwise. Here, instead of using $\hat{\mathbb{P}}(\tau|w;\mathbf{x})$, we adapt the leave-one-out estimation method by Li and Racine (2007) (see page 161 in Li and Racine (2007)) to define the likelihood function. Specifically, we apply $\hat{\mathbb{P}}_{-1}(\tau|w;\mathbf{x})$ in the likelihood function, which is in the same form as (45) but with $\mathcal{T}_{\mathbf{x}}$ replaced by $\mathcal{T}_{\mathbf{x}}\setminus\{(w,\tau)\}$. That is, the kernel probability density function for a sample labeled by w,τ, and $\mathbf{x}$ should exclude the sample itself in the summation. This method is available in the class "npcdensbw" of the R package "np". By carefully setting the input parameters in "npcdensbw", we obtain the bandwidths for mixed data types and obtain the estimated transition probabilities under KCDE.

With the estimated $\hat{\mathbb{P}}(\tau|w;\mathbf{x})$, we can calculate the probability that a new order arrives in period w since receiving the latest order as $\prod_{\tau=0}^{w-1}\hat{\mathbb{P}}(\tau+1|\tau;\mathbf{x})\cdot\hat{\mathbb{P}}(0|w;\mathbf{x})$. The probability of making consecutive orders can be calculated as $\sum_{w=0}^{T-1}\prod_{\tau=0}^{w-1}\hat{\mathbb{P}}(\tau+1|\tau;\mathbf{x})\cdot\hat{\mathbb{P}}(0|w;\mathbf{x})$.

B.2. Using RF To Estimate The Transition Probabilities and The Probability of Placing Consecutive Orders

RF is an ensemble learning method that builds a set of decision trees guiding classification results, which is available in many machine learning packages. We apply the class "RandomForestClassifier" in the python package "Scikit-learn" by Pedregosa et al. (2011) to predict probabilities. Recall that $\mathcal{L}=\{(0,\triangle),(0,M),(1,\triangle),(1,M),\cdots,(T-1,\triangle),(T-1,M)\}$ is a set that includes all the arcs to the terminal node Δ or M in network N. Let $\bar{q}^{RF}_{t,j}(\mathbf{x}^k_l)$, $j\in\{\triangle,M\}$, denote the predicted probability of the arc $(t,j)\in\mathcal{L}$. Given an order $\mathbf{x}^k_l$, the predicted probability $\bar{q}^{RF}_{t,j}(\mathbf{x}^k_l)$ of each class $t\in<T-1>$ and $j\in\{\Delta,M\}$ is the average of each single tree's probability of the class in the forest. A single tree's probability of a class is the fraction of training samples of the class over all the samples in the leaf where $\mathbf{x}^k_l$ locates (cf. Pedregosa et al. (2011)).

C. Introduction to AUC, Brier Score, and Calibration Curve

AUC is a popular performance metric to assess classifiers. The standard AUC is designed only for binary classification problems. Specifically, for a given feature-response samples $\{(\mathbf{x}^k, y^k)\}_{k\in[K]}$ where $y^k \in \{-1, 1\}$ is either -1 or 1, the AUC of a binary predictor $p: X \to [0,1]$ is defined as (Hand and Till 2001)

$$AUC(p(\cdot)) = \frac{\sum_{(\mathbf{x},y)\in\mathcal{D}^{-1}} \sum_{(\mathbf{x}',y')\in\mathcal{D}^1} \mathcal{I}[p(\mathbf{x}') > p(\mathbf{x})]}{|\mathcal{D}^{-1}| \cdot |\mathcal{D}^1|}, \tag{46}$$

where $p(\mathbf{x})$ represents the probability that the response is 1 conditioned on the feature is $\mathbf{x}$, $\mathcal{D}^i = \{(\mathbf{x}^k, y^k) | y^k = i, k \in [K]\}$ denotes the subsets of samples with their response as i for $i \in \{-1, 1\}$. From the definition, we can interpret AUC as the probability that a randomly selected positive sample shares higher predicted probability than a randomly selected negative sample. We call AUC for binary classification as AUC-binary.

Since predicting the choice probabilities in the consumer's sequential decision process is a multi-class classification problem, we need to discuss how to design AUC for a classification problem over multiple classes. We follow Hand and Till (2001) to generalize the AUC for the binary classification to the AUC for the multi-class classification. Specifically, for a I-class predictor $p(\mathbf{x}) = (p_1(\mathbf{x}), p_2(\mathbf{x}), \cdots, p_I(\mathbf{x})) : X \to \{(p_1, p_2, \cdots, p_I) | \sum_{i\in[I]} p_i = 1\}$ where $p_i(\mathbf{x})$ denotes the probability that the response is of i-th class conditioned on the feature is $\mathbf{x}$ for any $i \in [I]$, its AUC, denoted as $AUC(p(\cdot))$, is defined as

$$AUC(p(\cdot)) = \frac{1}{I(I-1)} \sum_{i=1}^{I} \sum_{j=1, j\neq i}^{I} AUC(i,j), \tag{47}$$

$$AUC(i,j) = \frac{\sum_{(\mathbf{x},y)\in\mathcal{D}^j} \sum_{(\mathbf{x}',y')\in\mathcal{D}^i} \mathcal{I}[p(\mathbf{x}') > p(\mathbf{x})]}{|\mathcal{D}^j| \cdot |\mathcal{D}^i|}, \tag{48}$$

where $\mathcal{D}^i = \{(\mathbf{x}^k, y^k) | y^k = i, k \in [K]\}$ denotes the subset of samples with their response as i for $i \in [I]$. We call this version of AUC for multi-class classification as AUC-Hand-Till.

The Brier Score of a multi-class predictor $p(\mathbf{x}) = (p_1(\mathbf{x}), p_2(\mathbf{x}), \cdots, p_I(\mathbf{x})) : X \to \{(p_1, p_2, \cdots, p_I) | \sum_{i\in[I]} p_i = 1\}$ is defined as

$$\sum_{k=1}^{K} \sum_{i=1}^{I} \|p(\mathbf{x}^k) - \mathbf{e}^k\|_2^2 \tag{49}$$

where $\mathbf{e}^k = (e_1^k, \cdots, e_I^k)$ and $e_i^k = 1$ if the kth sample is of the ith class, otherwise $e_i^k = 0$ for $i = 1, \cdots, I$.

As we have mentioned in Appendix B.2, both RF and MLE are predictors of the multi-class classification problem with label set as $\mathcal{L} = \{(0, \triangle), (0, M), (1, \triangle), (1, M), \cdots, (T-1, \triangle), (T-1, M)\}$, the above AUC-Hand-Till and Brier score can be used to assess their performance.

To plot a calibration curve under MLE or RF, we first sort all the orders according to their predicted probabilities of placing consecutive orders. Then, we divide the sorted orders into five even clusters. For each cluster, we compute its orders' average predicted probability of placing consecutive orders and the empirical percentage of the orders in the cluster that have a consecutive order. Each cluster corresponds to a point with its x coordinate equals the cluster's average predicted probability of placing consecutive orders, and its y coordinate equals the empirical percentage of the orders in the cluster that have a consecutive order. The calibration curve is formed by linking the points. Note that the nearer the calibration curve is to the diagonal line, the nearer the predictions are to their empirical counterparts.

D. Comparing Out-of-Sample Costs

We further compare the out-of-sample costs of MLE, KCDE, and RF. Table 7 reports the total out-of-sample costs of MLE and KCDE over the same 20 experiments described in Section 7.1.1. Again, we randomly sample 500 transaction paths for each experiment to reduce the computation time. The table shows that MLE outperforms KCDE based on their out-of-sample costs, which is consistent with the comparison results based on other performance metrics in Section 7.1.1. Table 8 shows that the out-of-sample costs under MLE and RF are still close to each other, with the relative cost gap (defined as $\frac{C_{RF}-C_{MLE}}{C_{MLE}} \times 100\%$) ranges from -0.056% to -0.011%.

Table 7: Out-of-sample costs of MLE and KCDE

$h(\times 10^{-3})$	Cost of KCDE	Cost of MLE	$GAP_{KCDE}(\%)$
0.262	10,174	10,164	0.098
0.861	10,169	10,155	0.138
1.775	10,145	10,117	0.277
2.788	10,109	10,065	0.437

Table 8: Out-of-sample costs of MLE and RF

$h(\times 10^{-3})$	Cost of RF	Cost of MLE	$GAP_{RF}(\%)$
0.262	298,120	298,152	-0.011
0.861	299,120	299,288	-0.056
1.775	299,674	299,786	-0.038
2.788	299,964	300,085	-0.040

E. Static Policy

The basic idea to get the optimal static threshold follows the idea of the Sample Average Approximation (SAA). Specifically, given a set of training samples $\{path^k\}_{k\in\mathcal{T}_1}$ ($\mathcal{T}_1$ is the index of training samples), the static threshold denoted by ST^* can be obtained by minimizing the empirical holding and order arrangement cost $f(ST;h,\{path^k\}_{k\in\mathcal{T}_1})$, which is defined as follows.

$$f(ST;h,\{path^k\}_{k\in\mathcal{T}_1}) = \sum_{k\in\mathcal{T}_1}\left\{\sum_{l=1}^{L_k-1}\mathcal{I}\{DT_l^k \le ST\}(DT_l^k\cdot h)+\mathcal{I}\{DT_l^k > ST\}(ST\cdot h+c)\right\} + \{ST\cdot h+c\}, \tag{50}$$

where the first term accounts for the total cost for the first L_k-1 consecutive orders whose subsequent choice is to make another order. For each order at position $l<L_k$, $DT_l^k \le ST$ indicates that a new order arrives by the holding threshold ST. According to the static policy, it only incurs a holding cost $DT_l^k\cdot h$ for the current order. In contrast, if $DT_l^k > ST$, a new order arrives after the holding threshold ST, in which case the current order has been sent to the 3PL. Therefore, not only the holding cost $ST\cdot h$ but also the order arrangement cost is incurred. The last term represents the cost incurred by the order at position L_k. By definition, no orders arrive for the maximal dwell time hence both holding and order arrangement cost is incurred.

With the optimal threshold ST^* obtained from the training samples, we can evaluate its performance on the testing samples $\{path^k\}_{k\in\mathcal{T}_2}$ where $\mathcal{T}_2$ is the indices of testing samples. The static policy is easy to implement and computationally efficient.

F. Estimating The Holding Cost Per Period h

We estimate the holding cost per period h as follows: (i) The retailer sets a target order arrangement fee reduction R. (ii) We determine the minimum threshold (number of periods) $\tau(R)$ to achieve the target. (iii) We find the maximum holding cost per period h such that the optimal static threshold is at least $\tau(R)$.

First, we can formulate a linear program to determine the threshold $\tau(R)$ to achieve the target of order arrange fee reduction.

$$\begin{aligned}\tau(R) = &\min_{\tau} \tau\\ &s.t.\ f(0;0,\{path^k\}_{k\in[K]}) - f(\tau;0,\{path^k\}_{k\in[K]}) \ge Rf(0;0,\{path^k\}_{k\in[K]})\\ &\quad\ \tau\in\langle 30\rangle,\end{aligned} \tag{51}$$

where R denotes the target order-arrangement cost reduction level and $0 \leq R \leq R_U = \frac{f(0;0,\{path^k\}_{k\in[K]})-f(30;0,\{path^k\}_{k\in[K]})}{f(0;0,\{path^k\}_{k\in[K]})}$. By setting h to 0, $f(\tau;0,\{path^k\}_{k\in[K]})$ represents the empirical order arrangement cost by holding the order for τ, which is decreasing in τ. Therefore, the problem can be solved via bisection search efficiently.

Given a threshold τ, we intend to find the maximal h such that the threshold is the optimal static threshold leading to the smallest total holding and order-arrangement cost, which serves as a soft constraint to ensure the timely delivery. Specifically, $h^*(\tau)$ can be estimated by solving the following linear program:

$$\begin{aligned} h^*(\tau) \in \arg\max_{h\geq 0} h \\ s.t.\ f(\tau;h,\{path^k\}_{k\in[K]}) \leq f(ST;h,\{path^k\}_{k\in[K]}), \forall ST \in \langle T\rangle, ST \neq \tau, \end{aligned} \tag{52}$$

where the cost function $f(ST;h,\{path^k\}_{k\in[K]})$ defined in (50) is linear in h. We start by setting τ to $\tau(R)$. If problem (52) with $\tau = \tau(R)$ is feasible, we set the holding cost per period to $h^*(\tau(R))$ for a given target R. If problem (52) with $\tau = \tau(R)$ becomes infeasible, we set h to $h^*(\tau^*)$ where τ^* is the smallest $\tau > \tau(R)$ such that the problem (52) is feasible. It is worthwhile to point out that we can always find such a τ^* as when $\tau^* = T$, $h = 0$ is feasible. With the estimation above, the optimal static policy of the system with the estimated h as the holding cost per period can achieve the target of order-arrangement reduction level R.

For the numerical experiments in Section 7, we use transaction paths with $T = 30$ minutes. We set $R = 0.85R_U, 0.65R_U, 0.45R_U$, and $0.25R_U$, where R_U =1.72%, which results in $h = 0.262 \times 10^{-3}, 0.861 \times 10^{-3}, 1.775 \times 10^{-3}$, and 2.788×10^{-3} respectively.

G. Tables

Table 9 Description of the "SKUs" table

Field	Data type	Description	Sample value
sku_ID	string	Unique identifier of a product	b4822497a5
type	int	1P or 3P SKU	1
brand_ID	string	Brand unique identification code	c840ce7809
attribute1	int	First key attribute of the category	3
attribute2	int	Second key attribute of the category	60
activate_date	string	The date at which the SKU is first introduced	2018/3/1
deactivate_date	string	The date at which the SKU is terminated	2018/3/1

Table 10 Description of the "users" table

Field	Data type	Description	Sample value
user_ID	string	User unique identification code	000000f736
user_level	int	User level	10
first_order_month	string	First month in which the customer placed an order on JD.com (format: yyyy-mm)	2017-07
plus	int	If user is with a PLUS membership	0
gender	string	User gender (estimated)	F
age	string	User age range (estimated)	26-35
marital_status	string	User marital status (estimated)	M
education	int	User education level (estimated)	3
purchase_power	int	User purchase power (estimated)	2
city_level	int	City level of user address	1

Table 11 Description of the "orders" table

Field	Data type	Description	Sample value
order_ID	string	Order unique identification code	3b76bfcd3b
user_ID	string	User unique identification code	3cde601074
sku_ID	string	SKU unique identification code	443fd601f0
order_date	string	Order date (format: yyyy-mm-dd)	2018/3/1
order_time	string	Specific time at which the order gets placed	2018-03-01 11:10:40
quantity	int	Number of units ordered	1
type	int	1P or 3P orders	1
promise	int	Expected delivery time (in days)	2
original_unit_price	float	Original list price	99.9
final_unit_price	float	Final purchase price	53.9
direct_discount_per_unit	float	Discount due to SKU direct discount	5
quantity_discount_per_unit	float	Discount due to purchase quantity	41
bundle_discount_per_unit	float	Discount due to bundle promotion	0
coupon_discount_per_unit	float	Discount due to customer coupon	0
gift_item	int	If the SKU is with gift promotion	0
dc_ori	int	Distribution center ID where the order is shipped from	29
dc_des	int	Destination address where the order is shipped to (represented by the closest distribution center)	29

Table 12 Description of the "clicks" table

Field	Data type	Description	Sample Value
sku_ID	string	SKU unique identification code	b4822497a5
user_ID	string	User unique identification code	94ff800585
request_time	string	The time at which the customer clicks the SKU item page	2018-03-01 23:57:53
channel	string	The click channel	wechart

Table 13 Summary of consecutive orders for each user level

user level	proportion of orders for each user level	number of non-consecutive orders	number of consecutive orders	**empirical percentage of the orders that have a consecutive order**	**average predicted probability of placing consecutive orders**
1 (Individual)	28.05%	120,227	809	0.72%	0.72%
2 (Individual)	30.87%	131,948	1,273	0.90%	0.90%
3 (Individual)	21.01%	89,654	1,014	1.08%	1.07%
4 (Individual)	18.96%	79,718	2,083	2.50%	2.49%
10 (Enterprise)	1.07%	3,347	1,291	27.84%	27.80%

Table 14 Summary of consecutive orders for each position

position	1	2	3	4
empirical percentage of the orders that have a consecutive order (%)	1.16	12.29	51.89	83.40
average predicted probability of placing consecutive orders (%)	1.16	12.22	51.42	82.74