Econometric Institute Report Series - EI2022-09

Customer-to-Customer Return Logistics: a New Way to Mitigate the Negative Impact of Online Returns

Ayse Sena Eruguz
Department of Operations Analytics, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands, a.s.eruguz@vu.nl,

Oktay Karabağ
Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands, karabag@ese.eur.nl,

Eline Tetteroo
TNO, 2597 AK The Hague, The Netherlands, elinetetteroo@gmail.com,

Carl van Heijst
It Goes Forward, Batavenpoort 36, 3991 JD, Houten, The Netherlands, carl@itgoesforward.com,

Wilco van den Heuvel, Rommert Dekker
Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands, wvandenheuvel@ese.eur.nl, rdekker@ese.eur.nl.

Abstract: Customer returns pose a big problem for retailers selling online due to high costs and CO_2 emissions. This paper introduces a new concept to handle online returns, the customer-to-customer (C2C) return logistics. The idea behind the C2C concept is to deliver returning items straight to the next customer, skipping retailers' warehouse. To incentivize customers to purchase C2C returning items, retailers can promote returning items on their webshop with a discount.We build the mathematical models behind the C2C concept to determine how much discount to offer, ensuring that enough customers are triggered to purchase C2C returning items and the expected total profit of the retailer is maximized. Our first model, the base model (BM), is a customer-based formulation of the problem and provides an easy-to-implement constant-discount-level policy. Our second model formulates the real-life problem as a Markov decision process (MDP). Since our MDP suffers from the curse of dimensionality, we resort to simulation optimization (SO) and reinforcement learning (RL) methods to obtain reasonably good solutions. We apply our methods using data collected from a Dutch fashion retailer. Furthermore, we provide extensive numerical experiments to claim generality. Our results indicate that the constant-discount-level policy obtained with the BM performs well in terms of expected profit compared to SO and RL. With the C2C concept, significant benefits can be achieved both in terms of expected profit and return rate. Even in cases where the cost-effectiveness of the C2C return program is not pronounced, the proportion of customer-to-warehouse returns to total demand gets lower. Hence, the system can be defined as more environmentally friendly. The C2C concept can help retailers in addressing the online return problem financially and adhering to the growing need for corporate social responsibility from the last decade.

Key words: reverse logistics; e-commerce; consumer returns; sustainability; Markov decision process

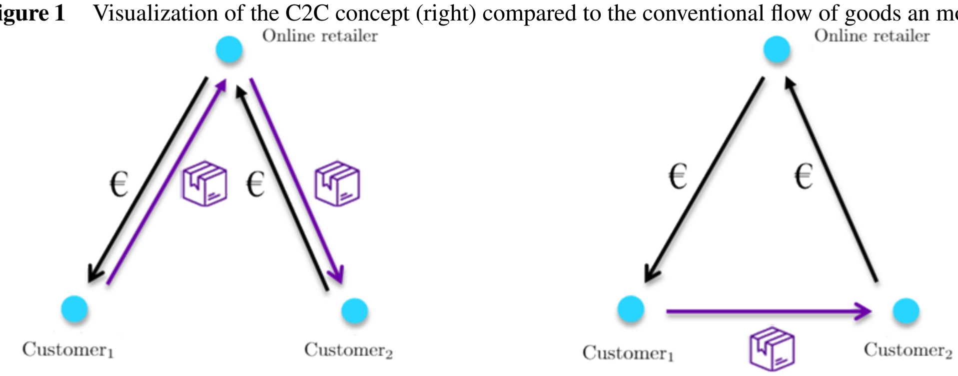
1 Introduction

Free product returns are an indispensable part of customer service in retail. As a result, large amounts of returns are sent back to warehouses of online retailers on a daily basis, with return rates varying from 5% up to 40% (NOS 2020, Dutch Broadcast Foundation), with similar figures found for the United States (data source: Statista 2018). The costs of return deliveries in the United States was estimated to be $350 billion in 2017 and $550 billion in 2020 (data source: Forbes 2019). Processing a returned product is a costly and time-consuming activity. It typically consists of collection, screening, and sometimes repairing (De Leeuw et al. 2016). One returned product is estimated to cost online retailers between €10.00 and €15.00 (NOS 2020), based on warehouse costs, labour costs, packaging costs, and transportation costs (both to and from the warehouse). Next to the requirement of additional warehouse logistics capabilities, returns also result in lower product availability and suboptimal reordering policies. Morever, during peak return periods, e.g. just after Christmas, retailers have to process vast amounts of returns. This burdens warehouse performance and incurs additional costs to manage the processing capacity. Clearly, the increasing amount of daily returns during the past decade suppresses the profitability of the online retail sector. Moreover, high return volumes have a negative environmental impact, because of the extra transport movements needed. Returns cause increased parcel express volumes and therefore CO_2 emissions. Besides that, each return requires new packaging material when being sold again.

Return rates turn out to be especially high in fashion e-commerce. Diggins et al. (2016) mention that return rates between 20% and 40% are common, with values of up to 74% (see Mostard and Teunter 2006). As customers typically return fashion products because of the wrong size or color, “a significant number of the returned products are in perfect condition for resale” (de Leeuw et al. 2016), which means there is no need for an immediate return to the warehouse.

Return volumes and associated costs show there is still no proper solution to the big return problem (online) retailers face on a daily basis. With this research, we propose a new concept as possible solution to the customer return problem, the ‘Customer-to-Customer’ (C2C) Return Logistics concept. In contrast to existing approaches where solutions to the return problem focus on preventing a return, we take a different approach and aim at minimizing the impact once a return happens. The main idea behind the concept is to send returned items directly to other customers demanding the same product instead of first being sent back to the retailer’s warehouse. An arriving customer is offered the possibility to get the product delivered directly from a returning customer, next to the conventional option to get the product delivered directly from the retailer’s warehouse. To compensate for a lower unboxing experience and a possible longer and variable delivery time, the option to buy a returned item is offered in exchange for a discount on the corresponding product. The customer returning the item could be incentivised to join the C2C return program in an economical way, for instance, by offering loyalty points, or in a moral way, by making him/her realize this is a 'greener' option compared to a direct return. The environment benefit comes from removing non-value adding transport movements and reducing packaging materials. The online retailer, which could also be an external party, has the managing and operating role in the process of the product flows through customers and offering of discounts. All transportation costs and cash flows are at the responsibility of the retailer. This means customers pay, as well as receive, money to and from the retailer, even though products are sent between customers. A visualisation of the C2C concept in its basic form is displayed in Figure 1. It is obvious from this figure that, besides cost savings, also CO_2 emission reduction can be achieved.

Figure 1 Visualization of the C2C concept (right) compared to the conventional flow of goods an money (left).



The concept resembles e-commerce sites selling used cloths, like Vinted (vinted.com). It is also comparable to retailers selling open-box items at a discount next to their new items. In contrast to Vinted, in our case (i) the retailer organizes the process and coordinates it with his direct sales and (ii) the clothes are not used, but can still be considered as good as new, hence serving a different market.

The contribution of our research is both practical and theoretical. From a practical point of view, this research introduces and analyzes a new concept, the C2C concept, which has the potential to decrease the physical and financial burden for retailers in handling returns. To the best of our knowledge, this concept has not yet been implemented by any retailer. In our research, we contacted multiple retailers who showed interest in the implementation of the concept. Co-authors of this article pursue these real-life implementations and ran promising experiments. Their progress can be followed on https://itgoesforward.com/. From a theoretical point of view, we provide the mathematical models which substantiate the C2C concept. First, we propose a stylized base model and provide some theoretical results regarding the structure of the profit function, the optimal discount level, and the profitability of the C2C return program. We propose a constant-discount-level policy for our problem by using our base model. Second, we formulate the real-life problem as an MDP. Because our MDP suffers from the curse of dimensionality, we resort to simulation optimization and reinforcement learning algorithms to find reasonably good solutions. As such, we are able to evaluate the financial prospects of the C2C concept, for various levels of product demand volumes, return rates, customer participation scenarios, and other model variables. The methods are applied to data collected from a

Dutch fashion retailer and consist of online sales and return data from May 2017 to May 2019, including 2.6 million data points. We find that the C2C concept could lead to higher profits compared to the conventional return program. In our case study, we evaluate best-case and worse case scenarios. Under the best case scenario, the C2C return program generates additional demand and we observe a profit increase of 34%. We show that the value of the C2C return program is non-significant under the worse-case scenario where C2C demand fully substitutes regular demand. Interestingly, in the cases of minor savings, the proportion of customer-to-warehouse returns to total demand decreases by 6–11%, which suggests that the C2C concept is still beneficial from an environmental point of view. In a more extensive set of experiments, in which we consider a full factorial combination of possible parameter values, this behavior is confirmed.

The remainder of this paper is organized as follows: Section 2 reviews the research most related to our paper. Section 3 elaborates on the C2C return program. Section 4 presents the base model under the conventional and the C2C return programs and provides theoretical results on the optimal discount level and the relation between the conventional and C2C return programs. Section 5 formulates the multi-customer model and provides theoretical results on the relation between the base and multi-customer models under certain assumptions. Section 6 presents solution methods to solve our MDP. Section 7 and 8 reports our comparative results for the performance of the algorithms and the value of the C2C program for our case study and numerical experiments, respectively. Finally, Section 8 concludes the paper with critical insights from the study and future extensions of our work.

2 Literature review

Although our contribution is rather unique, our research is part of the consumer returns management stream within closed-loop supply chains. Abdulla et al. (2019) give a recent overview of this area. Main research questions are under which conditions companies should allow returns, how much restocking fee to charge, how to collect returns, and what to do with them. In what follows, we summarize the most relevant papers related to our research.

Mollenkopf et al. (2007) find through empirical research that the return process is very important for customer loyalty. Their research highlights the importance to make the return process hassle-free. Thoughtfully crafted return policies positively influence consumers' perceptions and intention to repurchase. A similar conclusion is drawn by Griffis et al. (2012) who analyse purchase and return records of a moderately sized online retailer. Even though it is rare to observe charging re-stocking fees from customers in practice, this concept is well researched in the literature. An early seminal work by Shulman et al. (2009) propose an analytical model to examine how consumer purchase and return decisions are affected by a retailer's pricing and restocking fee decisions. Hjort and Lantz (2016) analyse the response of customers on the easiness to return. They find that free returns lead to increased order frequency, decreased average value of purchased items, increased probability of return, and increased average value of returned items. The results were based on experiments with a nordic e-commerce company (nelly.com). Abdulla et al. (2022) propose a cognitive process model to represent how consumers perceive, evaluate, and respond to return policies. Their model considers five return policy leniency levers: monetary, time, effort, scope, and exchange. They empirically test their model and show that the effect of monetary leniency dominates the effects of the other four levers.

Akçay et al. (2013) investigate the effect of selling returned products on the overall profit since they may cannibalize the sales of the new products. They analyse the pricing of the products under uncertain demand and include the option to sell returns at a fixed discount. We contribute to this stream by proposing a separate logistics channel to sell returned products and investigate how much discount we should give to make this system profitable.

Within e-commerce returns, fashion products have remarkably high return rates. Fashion products are often differentiated in size and fits and various sizes can be ordered by a single customer. As a consequence, many returns are of excellent quality, similar to cloths hanging in a shop and tried out by different customers. In many countries returned cloths may even be sold as new. A lot of fashion is seasonal, implying a short selling season. Thus, a fast processing of returns is required. De Leeuw et al. (2016) give a detailed analysis of the return process at fashion retailers. After inspection, cleaning, and sometimes repairing; products can be resold as new. Yet these activities have substantial costs. Diggins et al. (2016) provide a review on fashion returns management literature, quoting the special characteristics of fashion and concentrating very much at the marketing aspects of various return options. Difrancesco et al. (2018) develop a queuing model to optimize the return duration decision of a fashion retailer having a closed-loop supply chain with refurbishing. The authors investigate how to set the return duration, whether to refurbish returned products or sell them in the secondary market, so as to maximize the profit. They cross-check their analytical results with a data set coming from one of Europe's largest fashion online retailers, Zalando. Factors affecting the product returns in the apparel industry have been studied by several authors using either different data sets (Shang et al. 2017, 2019, Narang and Shankar 2019, Patel et al. 2021) or different focus groups (Lee 2015, Minnema et al. 2018). None of these papers consider selling returns with a discount and shipping them directly from customer-to-customer.

3 Customer-to-Customer Return Program

In this section, we describe the C2C return program and elaborate on the relation between program design and customer participation. First, we give a detailed explanation of the C2C return program. Second, we discuss incentives for customer participation in returning and purchasing items through the C2C return program. Third, we discuss how customer demand can be modelled. Finally, we present how to solve potential conflicts between customer and/or commercial entities and how to mitigate abuse.

The C2C return program consists of the following steps. First, the returning customer is asked to assess the condition of the item (e.g., as-good-as-new, defective, stained, etc.). If the item is assessed as as-good-as-new and the customer is willing to participate to the C2C return program, then the returning customer is asked to re-pack and keep the item during a so-called *time window for matching* (a few days). During this period, the item is offered on the webshop of the retailer as a discounted C2C item. Meanwhile, arriving customers are given two options on the webshop: (1) purchasing an item that is to be shipped from warehouse-to-customer (W2C) or (2) purchasing a discounted item to be shipped from customer-to-customer (C2C). When a C2C sale occurs, the QR code on item's repackaging material is linked to the purchasing customer's address and the returning customer is asked to hand-in the item within a so-called *time window for handing-in* (also a few days). The customer who purchased a C2C returning item is allowed to return the item to the warehouse if he/she is not satisfied with the purchase. However, the same item cannot be returned from C2C to enable a quality check and a package refurbishment by the retailer.

Time windows for matching and handing-in request additional efforts from returning customers. Returning customers should keep the item at most during the whole duration of the time window for matching. In addition, they should be able to hand-in the item on short notice either by going to a drop-off point or using a pick-up service. For a long time window for matching and a short time window for handing-in, returning customers' hassle increase, but also the likelihood of a C2C sale. Previous research on customer returns develop endogenous models by introducing a hassle cost of return (see, Xing et al. 2020, Shulman et al. 2009). In practice, it is difficult to quantify the hassle cost for returning customers. In order to substantiate the C2C initiative, Wiersma (2021) and Hsieh (2021) performed a rating-based conjoint analysis to collect and analyze the preferences of customers in the C2C program. According to their results, there are indications that the likelihood of returning customers to participate to the C2C return program is decreasing in the time window for matching and increasing in the time window for handing-in. In our paper, through numerical experiments, we assess the impact of these relations on the expected profit.

The C2C return program could lead to less transport movements, lower CO_2 emissions, and less packaging material compared to the current conventional return and purchase process of an item. Returning customers can be encouraged to participate by promoting the ecological benefits of the C2C return program. According to the recent questionnaire performed by Wiersma (2021) and Hsieh (2021), a significant share of respondents is willing to participate to the C2C return program under low or no monetary benefits. Self-esteem, altruism, and contributing to a better environment seems already enough to motivate individuals to participate to the C2C return program. In our paper, we ignore the effect of monetary benefits for returning customers. We consider that returning customers are solely triggered by moral incentives to perform a C2C return.

For purchasing customers, the C2C service differs from the conventional service in terms of 1) shipment time, 2) unboxing experience, 3) condition of the product, 4) additional discount offered. The C2C service offers the purchasing customer a potential inferior experience on the first three elements. Therefore, the retailer offers the C2C returning items at a discounted price. The discount level should be high enough such that enough customers are triggered to purchase C2C returning items. But, it should be as low as possible to maximize the expected profit. Wiersma (2021) and Hsieh (2021) show that the discount level is the most important attribute for purchasing customers (compared to delivery speed, type of product, and product value). In our paper, we express the C2C demand as an increasing function of the discount level and aim at optimizing the retailer's decision on the discount level.

We suppose that a certain fraction of existing customers would be attracted by the discount on the C2C returning items. In other words, the C2C return program would lead to demand substitution. However, there would also be new customers who are only attracted by the C2C returning items. These customers consider the conventional product price as too expensive and would not purchase an item otherwise. In our paper, we model the two extremes by formulating the best-case and the worst-case scenarios. In the best-case scenario, all C2C purchasing customers are new customers and the C2C return program increases the total demand. In the worst-case scenario, C2C returning items are seen as substitutes, hence the total demand remains the same. We assess the profitability of the C2C return program under these two extremes scenarios. We note that the C2C concept has not yet been used by retailers. But, most probably, the real situation would be somewhere in-between the best-case and the worst-case scenarios.

We note that only when the returning customer indicates the item is as-good-as-new, the returning customer is given the choice to return the item using the C2C return program. Damaged items will always be returned using the current conventional return program, i.e., from customer-to-warehouse (C2W). Indeed, the item has to be easily inspectable by the customers. Apparel, small home appliances, and products targeted to environmental focused customers are suitable for a C2C return.

In case a damaged item is returned to the retailer, the current conventional return program can lead to disputes between the returning customer, the third party logistics provider, and the retailer. Under the C2C return program, the purchasing customer is an additional actor, who is also a private individual. Currently, many commodity or services are exchanged between private individuals that are facilitated by a commercial party or platform. These companies make the platform economy of which Ebay, Vinted, Airbnb, Uber, and Lyft are best known. These platforms effectively deal with numerous disputes on a daily base. For the C2C return program, ItGoesForward would operate as a platform company in order to deal with disputes. Behaviour monitoring, identification/exclusion of service abusers, peer-to-peer reviews, and monetary compensation are just some of the tools platform economy companies can use to minimise, prevent, and, settle disputes between private individuals and disputes between a commercial party and a private individual. In this research, we exclude the associated costs in both conventional and C2C return programs, and we compare the two programs for as-good-as-new items.

In our mathematical models, we consider that all customers are trustworthy and do not abuse the system, for instance, by sending counterfeit items or performing pre-arranged C2C returns and purchases. In practice, the associated risks can be mitigated by using microchip tags on items (e.g., RFID and NFC), customer reputation systems, and identification/exclusion of service abusers by behaviour monitoring. The elements of community building and customer reviews are factors that can foster the adoption of the C2C return program.

4 Base Model

In this section, we define our problem and the base model under the conventional and C2C return programs. The base model employs a customer-based approach by considering a single customer initially served from the warehouse.

4.1 Conventional return program

We consider an online retailer that sells an item in a webshop at price $P > 0$. Upon purchase, the item is sent from warehouse to customer (W2C), incurring handling and shipment cost S^{W2C}. Under the conventional return program, the customer is allowed to return the item within T^{R} periods after delivery. Upon a return from customer to warehouse (C2W), shipment and handling cost S^{C2W} is incurred. We assume that the customer is fully refunded. We model the customer's return decisions as time-varying Bernoulli trials. Let u_i^{R} be the probability of return i periods after delivery with $i \in \{1, 2, \ldots, T^{\mathrm{R}}\}$. The probability of a C2W return is

$$p^{\mathrm{R}} = \sum_{i=1}^{T^{\mathrm{R}}} u_i^{\mathrm{R}} \prod_{j=1}^{i-1} \left(1 - u_j^{\mathrm{R}}\right) = 1 - \prod_{i=1}^{T^{\mathrm{R}}} \left(1 - u_i^{\mathrm{R}}\right),$$

where the equality follows from $\prod_{i=1}^{T^{\mathrm{R}}} (1 - u_i^{\mathrm{R}})$ being the probability of no return. Under the natural assumption that $u_i^{\mathrm{R}} < 1$ for $i \in \{1, 2, \ldots, T^{\mathrm{R}}\}$, we have $p^{\mathrm{R}} < 1$. Let $R^{W2C} = P - S^{\mathrm{W2C}}$ be the revenue generated by a W2C delivery and $C^{C2W} = P + S^{\mathrm{C2W}}$ be the cost incurred due to a C2W return. Under the conventional return program, the expected profit of the retailer from a single customer is

$$\mathbb{E}\left[\Pi\right] = R^{W2C} - C^{C2W} p^{\mathrm{R}}. \tag{1}$$

In this profit function, ordering, purchasing, inventory holding related costs are disregarded since they are not impacted by the return program.

4.2 C2C return program

Under the C2C return program, returning customers can either choose a conventional C2W return or a C2C return, within T^{R} periods after delivery. In this case, we distinguish two cases: (1) best-case in which C2C return program potentially brings an additional customer and (2) worst-case in which a C2C returning item is considered as a substitute by an arriving customer.

4.2.1 Return process.

Consider a customer whose item was delivered i periods ago from W2C with $i \in \{1,2,\ldots,T^{\mathrm{R}}\}$. If the item has not been returned yet, the customer can (1) announce a conventional C2W return with probability u_i^{C2W}, (2) announce a C2C return with probability u_i^{C2C}, (3) keep the item one more period with probability $1-u_i^{\mathrm{C2W}}-u_i^{\mathrm{C2C}}$. The probability of a C2C return is

$$p^{\mathrm{C2C}} = \sum_{i=1}^{T^{\mathrm{R}}} u_i^{C2C} \prod_{j=1}^{i-1} \left(1-u_j^{C2C}-u_j^{C2W}\right).$$

Similarly, the probability of a conventional C2W return is

$$p^{\mathrm{C2W}} = \sum_{i=1}^{T^{\mathrm{R}}} u_i^{C2W} \prod_{j=1}^{i-1} \left(1-u_j^{C2C}-u_j^{C2W}\right).$$

As in the conventional return program, we make the natural assumption that $u_i^{\mathrm{C2W}}+u_i^{\mathrm{C2C}}<1$ for $i \in \{1,2,\ldots,T^{\mathrm{R}}\}$, implying that $p^{\mathrm{C2C}}+p^{\mathrm{C2W}}<1$. Furthermore, a C2W return incurs cost C^{C2W}.

4.2.2 Matching process.

Following the announcement of a C2C return, the returning item is available in the webshop for sale. The retailer offers a discount $a \in [0,1]$ on the selling price P for this item. Let the probability of selling a C2C returning item in a given period $q(a)$ be a function of the discount level a. The probability that a C2C returning item is sold within the time window for matching T^{M} is

$$p^{\mathrm{M}}(a) = 1-(1-q(a))^{T^{\mathrm{M}}}.$$

If the C2C returning item is sold at discount level a, the resulting revenue is $R^{\mathrm{C2C}}(a)=(1-a)P$. If the returning item cannot be sold within T^{M} periods, a conventional C2W return is performed at cost C^{C2W}.

4.2.3 Hand-in process.

Following the matching of a C2C return to a C2C demand, the returning customer should hand in the item within the time window for handing-in T^{H}. The C2C returning customer hands in the item i periods after matching with probability u_i^{HI} for $i \in \{1,2,\ldots,T^{\mathrm{H}}\}$. Hence, the probability of hand-in within T^{H} periods after matching is

$$p^{\mathrm{HI}} = \sum_{i=1}^{T^{\mathrm{H}}} u_i^{\mathrm{HI}} \prod_{j=1}^{i-1} \left(1-u_j^{\mathrm{HI}}\right).$$

If the C2C returning customer hands in the item within T^{H} periods, the item is delivered from C2C and a C2C shipment cost S^{C2C} is incurred. The corresponding hand-in cost is $C^{\mathrm{HI}}=S^{\mathrm{C2C}}$. Otherwise, a new item is sent from W2C to satisfy the C2C demand, incurring handling and shipment cost S^{W2C}. Hence, the cost associated with a late hand-in is $C^{\mathrm{LHI}}=S^{\mathrm{C2W}}+S^{\mathrm{W2C}}$. When the C2C return program is launched, ensuring full refunds would encourage more customers to join the C2C return program. Thus, in our model, we consider full refunds for both late and on-time hand-ins.

4.2.4 Finalization process.

The customer who purchased the C2C returning item is allowed to return it within T^{R} periods. However, another C2C return is not allowed since a quality check and a package refurbishment should be performed by the retailer. We model this final return process as in the conventional return program. Let u_i^{FR} be the probability of return i periods after delivery with $t \in \{1,2,\ldots,T^{\mathrm{R}}\}$. The item is returned with probability

$$p^{\mathrm{FR}} = \sum_{i=1}^{T^{\mathrm{R}}} u_i^{\mathrm{FR}} \prod_{j=1}^{i-1} \left(1 - u_j^{\mathrm{FR}}\right),$$

where we make the natural assumption that $u_i^{\mathrm{FR}} < 1$ for $i \in \{1,2,\ldots,T^{\mathrm{R}}\}$, implying that $p^{\mathrm{FR}} < 1$. The cost associated with the return of a C2C customer who purchased an item at discount level a is $C^{\mathrm{C2C}}(a) = (1-a)P + S^{\mathrm{C2W}}$.

4.2.5 Expected profit.

We express the expected profit for the best case (i.e., the number of customers potentially increases) and the worse case (i.e., the number of customers remains the same).

Given discount level a, the expected best-case profit of the retailer $\mathbb{E}\left[\Pi_{\mathrm{B}}(a)\right]$ can be written as

$$\begin{aligned}\mathbb{E}\left[\Pi_{\mathrm{B}}(a)\right] =& R^{W2C} - C^{C2W}\left[p^{\mathrm{C2W}} + p^{\mathrm{C2C}}\left(1 - p^{\mathrm{M}}(a)\right)\right] + R^{\mathrm{C2C}}(a)p^{\mathrm{C2C}}p^{\mathrm{M}}(a)\\ &- \left[C^{\mathrm{HI}}p^{\mathrm{HI}} + C^{\mathrm{LHI}}\left(1 - p^{\mathrm{HI}}\right) + P\right]p^{\mathrm{C2C}}p^{\mathrm{M}}(a) - C^{\mathrm{C2C}}(a)p^{\mathrm{C2C}}p^{\mathrm{M}}(a)p^{\mathrm{FR}}. \qquad (2)\end{aligned}$$

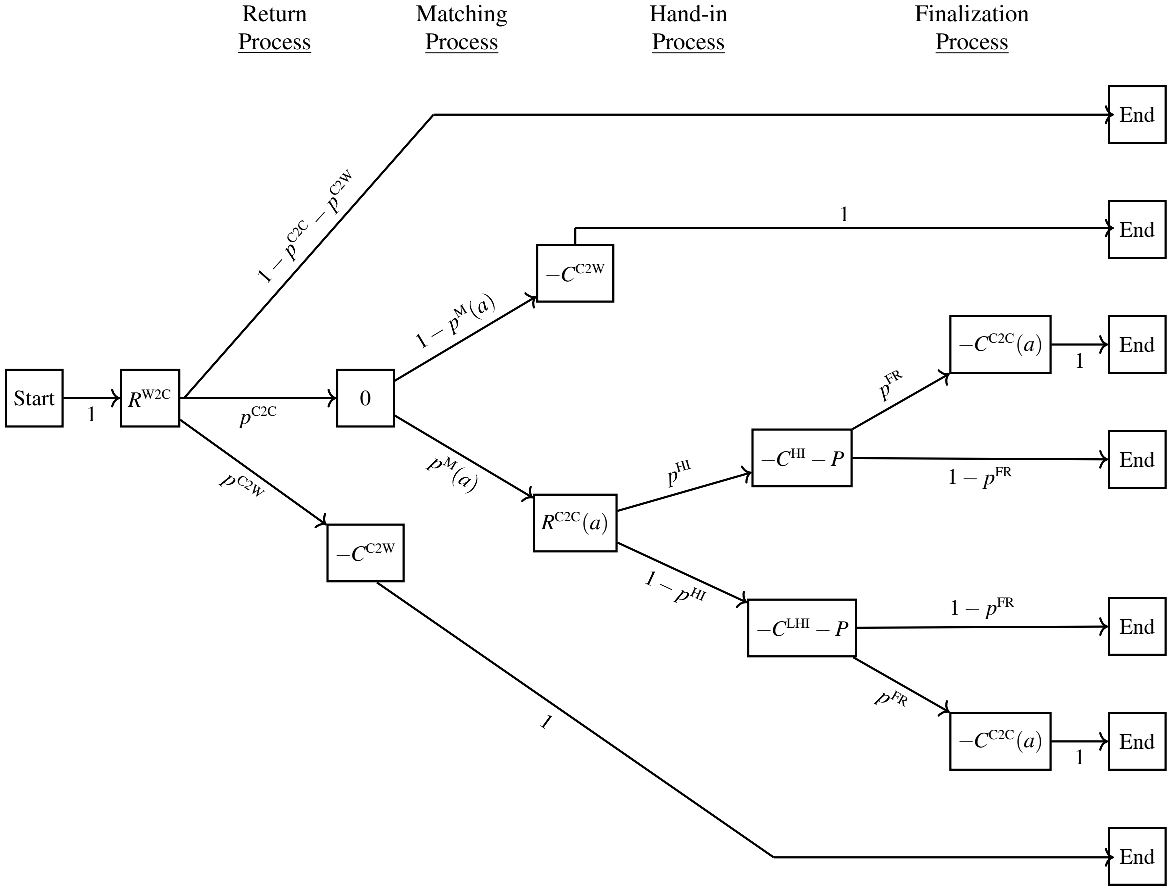
The first component of (2) is the revenue generated by a W2C delivery. The second component is the expected cost associated with a C2W return which is incurred if the customer (i) requests a C2W return or (ii) announces a C2C return but the item cannot be sold within T^{M} periods. The third component represents the expected revenue generated from a C2C sale. The fourth component represents the expected costs associated with a late or on-time hand-in of a C2C returning item and the refund to the returning customer. The fifth component is the expected cost associated with the return of a C2C demand. Figure 2 summarizes the corresponding processes, costs, and revenues under the C2C return program. The problem of the retailer is to determine an optimal discount level a^* that maximizes the expected profit function given in equation (2).

Given discount level a, an arriving customer purchases a C2C returning item with probability $p^{\mathrm{C2C}}p^{\mathrm{M}}(a)$. Under the best case scenario, the total demand increases by the ratio $p^{\mathrm{C2C}}p^{\mathrm{M}}(a)$. Under the worst-case scenario, the total demand remains the same. Hence, the expected worst-case profit of the retailer $\mathbb{E}\left[\Pi_{\mathrm{W}}(a)\right]$ can be expressed by

$$\mathbb{E}\left[\Pi_{\mathrm{W}}(a)\right] = \frac{\mathbb{E}\left[\Pi_{\mathrm{B}}(a)\right]}{1 + p^{\mathrm{C2C}}p^{\mathrm{M}}(a)}. \qquad (3)$$

Note that these profit functions are built based on the expected profit to be generated by an arriving customer and not based on an item in stock. By assuming ample stock in the warehouse, we make the connection to the multi-customer case in Section 5.

Figure 2 Processes, costs, and revenues under the C2C return program



4.3 Theoretical Results

In this section, we provide our structural results on the relation between the conventional and C2C return programs, and the optimal discount level under certain assumptions. Proofs of the theorems are provided in the Appendix.

ASSUMPTION 1. *The probability of a return in the conventional return program is equal to the probability of a return in the C2C return program, i.e.,* $p^R = p^{C2C} + p^{C2W}$.

This assumption seems reasonable as it is likely that a customer first makes a decision on returning an item, and given this decision, he/she will decide on which return program (conventional or C2C) to use. The following parameters play an important role in the structural results:

$$\begin{aligned} \phi_B &= C^{\mathrm{HI}} p^{\mathrm{HI}} + C^{\mathrm{LHI}} \left(1 - p^{\mathrm{HI}}\right) - (1 - p^{\mathrm{FR}}) S^{\mathrm{C2}W}, \\ \phi_W &= \phi_B + \mathbb{E}(\Pi). \end{aligned}$$

The term ϕ_B consists of the cost of handing-in ($C^{\mathrm{HI}} p^{\mathrm{HI}} + C^{\mathrm{LHI}} (1 - p^{\mathrm{HI}})$) and the benefit of eliminating the shipment from C2W (which is $(1 - p^{\mathrm{FR}}) S^{\mathrm{C2W}}$ since it only happens if the C2C purchase is not returned).

Thus, ϕ_B and ϕ_W can be interpreted as the difference in operational cost when using the C2C return program instead of the conventional for the best-case and worse-case scenarios, respectively.

THEOREM 1. *Under Assumption 1 and for a discount level $a \in [0,1]$, the C2C return program is more profitable than the conventional one in case of the best-case scenario if and only if*

$$P(1-a)\left(1-p^{FR}\right) \geq \phi_B, \tag{4}$$

and in case of the worst-case scenario if and only if

$$P(1-a)\left(1-p^{FR}\right) \geq \phi_W. \tag{5}$$

It is interesting to note that the profitability of the C2C return program in comparison to the conventional one is independent of the probability $p^{\mathrm{M}}(a)$, that is, it is independent of the functional relation between the C2C demand and the discount level. In the best-case scenario, one needs to check whether the discount level a is such that the revenue from selling the item to the C2C customer (i.e., $P(1-a)(1-p^{\mathrm{FR}})$) exceeds the difference in operational cost. As a consequence, in case $\phi_B < 0$ the C2C return program is always (i.e., no matter which discount level a is chosen) more profitable, while the opposite is true in case $\phi_B > P(1-p^{\mathrm{FR}})$ (i.e., even the highest revenue corresponding to discount level $a = 0$ cannot compensate for the cost). A similar reasoning holds for the worst-case scenario, but the revenue also needs to make up for $\mathbb{E}(\Pi)$, as under this scenario a conventional customer is substituted by a C2C customer. To summarize the discussion, there exists a discount level $a \in [0,1]$ for which the C2C return program is more profitable than the conventional return program in case of the best-case (resp. worst-case) scenario if and only if $P(1-p^{\mathrm{FR}}) \geq \phi_B$ (resp. $P(1-p^{\mathrm{FR}}) \geq \phi_W$).

Finally, conditions (4) and (5) provide an upper bound on the optimal discount level a, assuming that a firm only prefers to offer the C2C return program if it is more profitable than the conventional program. To get tractable expressions for the optimal discount level, we need the following assumption.

ASSUMPTION 2. *The probability of selling a C2C returning item $q(a)$ is a linear function of discount level a, i.e., $q(a) = q_0 + a(1-q_0)$ where q_0 is the probability of selling the C2C returning at its original price P and* $0 \leq q_0 < 1$.

While it is possible to specify any functional form for $q(a)$, we consider a simple linear function for the ease of analysis of our base model (see, e.g., Adida and Perakis 2009, Cohen et al. 2021, for other studies that make a similar assumption).

THEOREM 2. *The best-case profit function $E[\Pi_B(a)]$ is unimodal on the interval* $[0,1]$.

Theorem 2 immediately provides a method to numerically find the optimal best-case discount level a_B^*. Namely, it is well known that golden-section search is a method that can efficiently find the optimum of a unimodal function at any desired precision. Although we conjecture that $E[\Pi_{\text{W}}(a)]$ is unimodal as well, we are unable to prove this formally. However, the next property turns out to be useful in the sense that it reduces the search space for finding the optimal discount level in the worst case model, besides the fact that it is of interest on its own.

THEOREM 3. *The optimal best-case discount level a_B^* equals at most the optimal worst-case discount level a_W^*, i.e., $a_W^* \leq a_B^*$.*

For certain parameter ranges (see Appendix, Theorem EC.1), we have an analytical expression for the optimal discount levels, although in many cases one needs to solve a high degree polynomial equation, for which no closed-form solution exists. Ways to still find a close-to-optimal discount level are to (i) use a brute force method, for example, by simply enumerating over a finite set of discount levels, for instance, $a = 0\%, 1\%, \ldots, 100\%$, or (ii) use golden-section search which leads to a locally but not necessary globally optimal solution (as $E[\Pi_{\text{W}}(a)]$ is conjectured to be unimodal).

5 Multi-Customer Model

In this section, we present the multi-customer model under the conventional and C2C return programs. The multi-customer model extends the base model to multiple arriving and returning customers during a selling season of T periods. Under the C2C return program, the main difference compared to the base model is the explicit modeling of C2C return-demand matching process and the influence of discount levels on both C2C and W2C demands.

5.1 Conventional return program

We assume that customers arrive according to a Poisson process with rate λ_t during $[t, t+1)$ for $t \in \{0, 1, \ldots, T-1\}$. The demand size of each customer is one unit. We assume ample stock in the warehouse. Demands occurring during $[t, t+1)$ are satisfied at the end of time t, generating revenue R^{W2C} per customer. Let $\boldsymbol{x}_t = (x_{t,i} \mid i \in \{1, \ldots, T^{\text{R}}\})$ be a T^{R} - dimensional vector of integers, where $x_{t,i}$ represents the number of customers satisfied i periods ago and not returned by time t. For items purchased at time $t-i$, the number of C2W returns during $[t, t+1)$ follows a binomial distribution with parameters $x_{t,i}$ and u_i^{R} with $i \in \{1, \ldots, T^{\text{R}}\}$. Each C2W return incurs cost C^{C2W}. Due to the relation between Bernoulli trials and the binomial distribution, this model scales up the base model given in Section 4.1 in a stationary demand, infinite horizon setting (see Theorem 4).

5.2 C2C return program

We formulate the multi-customer problem under the C2C return program as a finite-horizon Markov Decision Process (MDP). In what follows, we introduce the elements of our MDP problem, namely, state information, actions, exogenous information, transition function, and objective function.

The sequence of events is as follows. At the beginning of time t, state $\boldsymbol{s}_t$ is observed and a decision a_t is made. During $[t, t+1)$, exogenous information $\boldsymbol{w}_t$ is faced. At the end of time t, costs and revenues are evaluated.

5.2.1 State information.

At time t, state information $\boldsymbol{s}_t = (\boldsymbol{x}_t, \boldsymbol{y}_t)$ is composed of

(i) $\boldsymbol{x}_t = (x_{t,i} \mid i \in \{1, \ldots, T^{\mathrm{R}}\})$ where $x_{t,i}$ represents the number of W2C demands which are satisfied i periods ago and not returned by time t,

(ii) $\boldsymbol{y}_t = (y_{t,i} \mid i \in \{1, \ldots, T^{\mathrm{M}}\})$ where $y_{t,i}$ represents the number of C2C returns which are announced i periods ago and not matched to a C2C demand by time t.

5.2.2 Actions.

Action $a_t \in [0, 1]$ is the discount level applied to C2C demands during time $[t, t+1)$.

5.2.3 Exogenous information.

Exogenous information $\boldsymbol{w}_t = (d_t^{\mathrm{W2C}}, d_t^{\mathrm{C2C}}, \boldsymbol{r}_t^{\mathrm{C2W}}, \boldsymbol{r}_t^{\mathrm{C2C}})$ known at the end of time t is composed of

(i) d_t^{W2C}, the number of newly arrived W2C demands,

(ii) d_t^{C2C}, the number of newly arrived C2C demands,

(iii) $\boldsymbol{r}_t^{\mathrm{C2W}} = \left(r_{t,i}^{\mathrm{C2W}} \mid i \in \{1, \ldots, T^{\mathrm{R}}\}\right)$ where $r_{t,i}^{\mathrm{C2W}}$ is the number of newly announced C2W returns of items purchased at time $t - i$,

(iv) $\boldsymbol{r}_t^{\mathrm{C2C}} = \left(r_{t,i}^{\mathrm{C2C}} \mid i \in \{1, \ldots, T^{\mathrm{R}}\}\right)$ where $r_{t,i}^{\mathrm{C2C}}$ is the number of newly announced C2C returns of items purchased at time $t - i$.

Note that d_t^{W2C} and d_t^{C2C} would depend on the discount level a_t and current time t. Typically, d_t^{W2C} would decrease and d_t^{C2C} would increase with the discount level a_t in time t. We model the two extremes by the following best-case and worst-case demand scenarios.

In the best-case demand scenario, W2C demand remains the same and total customer demand increases by C2C demand. We model this scenario as follows. We assume that W2C demand is Poisson with rate λ_t during $[t, t+1)$ where λ_t. C2C demand d_t^{C2C} follows a binomial distribution with parameters $\sum_{i=1}^{T^{\mathrm{M}}} y_{t,i}$ and $q(a_t)$ during $[t, t+1)$.

In the worst-case demand scenario, total customer demand remains the same, i.e., W2C demand decreases by C2C demand. Both W2C and C2C demands depend on the current discount level. We model this scenario as follows. Customers arrive according to a Poisson process with rate λ_t during $[t,t+1)$ where λ_t. An arriving customer is willing to purchase a C2C returning item with probability $q(a_t)$. The number of customers who are willing to purchase a C2C returning item $\tilde{d}_t^{\text{C2C}}$ follows a binomial distribution with parameters d_t and $q(a_t)$ where d_t is the number of newly arrived customers during $[t,t+1)$. When a new customer arrives, if there are no C2C returns available for purchase, purchasing a C2C returning item is not an option for the arriving customer. Therefore, we have $d_t^{\text{C2C}} = \min\{\tilde{d}_t^{\text{C2C}}, \sum_{i=1}^{T^{\text{M}}} y_{t,i}\}$ and $d_t^{\text{W2C}} = d_t - d_t^{\text{C2C}}$.

The probability of a customer return depends on time elapsed since delivery. We assume that newly announced C2W and C2C returns $(r_{t,i}^{\text{C2W}}, r_{t,i}^{\text{C2C}})$ follow a multinomial distribution with parameters $x_{t,i}$ and $(u_i^{\text{C2W}}, u_i^{\text{C2C}})$ during $[t,t+1)$.

5.2.4 Transition function.

The transition function $S(.)$ defines the transition from state $\boldsymbol{s}_t$ to $\boldsymbol{s}_{t+1} = S(\boldsymbol{s}_t, a_t, \boldsymbol{w}_t)$ after taking action a_t and facing exogenous information $\boldsymbol{w}_t$. We split the transition function into $S^1(.)$ and $S^2(.)$ where $\boldsymbol{x}_{t+1} = S^1(\boldsymbol{x}_t, a_t, d_t^{\text{W2C}}, \boldsymbol{r}_t^{\text{C2W}}, \boldsymbol{r}_t^{\text{C2C}})$ and $\boldsymbol{y}_{t+1} = S^2(\boldsymbol{y}_t, a_t, d_t^{\text{C2C}}, \boldsymbol{r}_t^{\text{C2C}})$. Functions $S^1(.)$ and $S^2(.)$ generate T^{R} and T^{M} -dimensional vectors, respectively.

At the transition from time t to $t+1$, state information is advanced by 1 period. During $[t,t+1)$, the number of W2C demand is d_t^{W2C} and the number of returns is $\boldsymbol{r}_t^{\text{C2W}} + \boldsymbol{r}_t^{\text{C2C}}$. Transition function $S^1(.)$ can be defined as

$$\left(S^1(\boldsymbol{x}_t, a_t, d_t, d_t^{\text{C2C}}, \boldsymbol{r}_t)\right)_i = \begin{cases} d_t^{\text{W2C}} & \text{for } i = 1, \\ x_{t,i-1} - r_{t,i-1}^{\text{C2W}} - r_{t,i-1}^{\text{C2C}} & \text{for } i = 2,3,\ldots,T^{\text{R}}, \end{cases}$$

where $(.)_i$ is the i^{th} element of the vector inside the brackets.

During $[t,t+1)$, the number of C2C returns is $\sum_{j=1}^{T^{\text{R}}} r_{t,j}^{\text{C2C}}$ and the number of C2C demand is d_t^{C2C}. At the end of time t, C2C returns are matched to C2C demands according to the first-in first-out (FIFO) rule. That is, the matching starts from the oldest C2C returning items. Returning items announced T^{M} periods ago are matched first, those announced $T^{\text{M}}-1$ periods ago are matched second, etc., until the C2C demand is fully satisfied. Let $(x)^+ = \max(0,x)$. Define $d_{t,i}^{\text{C2C}} = \left(d_t^{\text{C2C}} - \sum_{j=i}^{T^{\text{M}}} y_{t,j}\right)^+$ as the remaining number of C2C demands after matching customers who announced a C2C return i periods ago and onward. Transition function $S^2(.)$ can be defined as

$$\left(S^2(\boldsymbol{y}_t, a_t, d_t^{\text{C2C}}, \boldsymbol{r}_t^{\text{C2C}})\right)_i = \begin{cases} \sum_{j=1}^{T^{\text{R}}} r_{t,j}^{\text{C2C}} & \text{for } i = 1, \\ \left(y_{t,i-1} - d_{t,i}^{\text{C2C}}\right)^+ & \text{for } i = 2,3,\ldots,T^{\text{M}}. \end{cases}$$

5.2.5 Objective function.

We consider costs, revenues, hand-in, and return probabilities as in the base model. The expected immediate profit at the end of time t, given state $\boldsymbol{s}_t$, action a_t, and exogenous information $\boldsymbol{w}_t$ is

$$\begin{aligned}\mathbb{E}\left[R(\boldsymbol{s}_t, a_t, \boldsymbol{w}_t)\right] =& R^{\text{W2C}} d_t^{\text{W2C}} - C^{\text{C2W}} \left[\sum_{i=1}^{T^{\text{R}}} r_{t,i}^{\text{C2W}} + \left(y_{t,T^{\text{M}}} - d_t^{\text{C2C}}\right)^+\right] + R^{\text{C2C}}(a_t) d_t^{\text{C2C}} \\ &- \left(C^{\text{HI}} p^{\text{HI}} + C^{\text{LHI}}(1 - p^{\text{HI}}) + P\right) d_t^{\text{C2C}} - C^{\text{C2C}}(a_t) p^{\text{FR}} d_t^{\text{C2C}}. \end{aligned} \quad (6)$$

Equation (6) follows the same reasoning as (2).

Let $V_t(\boldsymbol{s}_t)$ be the maximum expected total profit in state $\boldsymbol{s}_t$ following the optimal policy from time t onward. The optimal discount levels at time $t = 0, 1, \ldots, T-1$ can be obtained by

$$a_t^* = \underset{a_t \in [0,1]}{\arg\max} \{\mathbb{E}_{\boldsymbol{w}_t} [\mathbb{E} [R(\boldsymbol{s}_t, a_t, \boldsymbol{w}_t)] + V_{t+1}(\boldsymbol{s}_{t+1})]\},$$

where $V_T(\boldsymbol{s}) = 0$ for all states $\boldsymbol{s}$.

5.3 Theoretical Results

In this section, we provide our theoretical results on the relation between the base and multi-customer models under certain assumptions.

ASSUMPTION 3. *W2C demand follows a Poisson distribution with constant rate* λ.

THEOREM 4. *Let* $\mathbb{E}[\hat{\Pi}]$ *be the long-run average profit of the conventional multi-customer system defined in Section 5.1. Under Assumption 3, we have*

$$\mathbb{E}[\hat{\Pi}] = \lambda \mathbb{E}[\Pi].$$

Theorem 4 implies that the base and multi-customer models are equivalent to each other in a stationary-demand, infinite horizon setting. More specifically, when the expected profit of the base model is scaled up by the demand rate λ, we obtain the long-run average profit of the multi-customer model.

Let a constant-discount-level policy with parameter $a \in [0,1]$ be a policy with $a_t = a$ for all $t = \{1, 2, \ldots, T\}$.

THEOREM 5. *Let* $\mathbb{E}[\hat{\Pi}(a)]$ *be the long-run average profit of the C2C multi-customer system under a constant-discount-level policy with parameter a as defined in Section 5.2. Assume* $T^M = 1$*. Under Assumptions 3 and best-case scenario, we have*

$$\mathbb{E}[\hat{\Pi}(a)] = \lambda \mathbb{E}[\Pi_B(a)].$$

Theorem 5 implies that if $T^{\mathrm{M}} = 1$, then the base and multi-customer models are equivalent to each other in an infinite horizon setting and best-case scenario. We note that this result does not hold for $T^{\mathrm{M}} > 1$ due to the existence of the FIFO rule in the matching process. In addition, this result does not hold in case of worst-case demand since W2C demand would not follow Poisson distribution.

6 Solution Methods

The base model (BM) introduced in Section 4.2 simplifies the real-life problem under the C2C return program. The multi-customer model presented in Section 5.2 is more comprehensive, which is an MDP with unbounded state space and continuous action space. For the multi-customer model, finding an optimal policy via exact algorithms is computationally intractable. Thus, we propose heuristic approaches to find reasonably good solutions.

First, we consider a constant-discount-level policy with parameter a^{BM}, where a^{BM} is the optimal discount level for the base model under the best/worst-case demand scenarios, found by using Theorem 2 and 3, respectively. Note that using Theorem 2, we obtain the optimal constant-discount-level policy for $T^{\text{M}} = 1$ in a stationary demand, infinite horizon, best-case demand setting (Theorem 5).

Second, we determine a constant-discount-level policy by a simple simulation optimization (SO) procedure. Given resolution N, we evaluate the performance of a finite set of constant-discount-level policies with parameter $a \in \{0, 1/N, 2/N, ..., N-1/N\}$ by simulation, using common random variables. As such, we determine the discount level a^{SO} that provides the greatest expected total profit, where we also make statistical comparisons.

Third, we implement the reinforcement learning (RL) method introduced by Kearns et al. (2002). The algorithm we implemented can be found in Figure 1, page 198 of Kearns et al. (2002). This algorithm is designed for finding near optimal solutions to MDPs with infinitely large state spaces. It is based on the idea of *sparse sampling*, leading to a non-stationary stochastic policy. Given any state $\boldsymbol{s}_t$ at time t, the algorithm uses a *simulator* of the MDP to draw samples for many state-action pairs, and uses these samples to compute a good action from $\boldsymbol{s}_t$, which is then executed. More precisely, for state $\boldsymbol{s}_t$ at time t, a finite subset of actions A are considered and a randomly sampled look-ahead tree of depth H and sample size C is constructed. Using this look-ahead tree, we formulate a sub-MDP. The optimal action for this sub-MDP is obtained by *dynamic programming*. The complexity of the per-state computations (i.e., the number of simulated transitions for the development of the look-ahead tree) is $O\left((|A|C)^H\right)$. We note that for our problem, there will be no guarantee that the sub-MDP contains enough information to compute a near-optimal action from state $\boldsymbol{s}_t$. The number of calls to the simulator required to obtain a near-optimal solution is often extremely large (Péret and Garcia 2004). In exchange for this limitation, the running time of the algorithm has no dependence on the number of states. We use the solution obtained by this algorithm as a benchmark solution to assess the performance of the constant-discount-level policies obtained by the base model and simulation optimization for real-life cases introduced in Section 7. We refer to Kearns et al. (2002) for more details about this algorithm.

7 Case Study

In this section, we assess the value of the C2C return program at a fashion retailer in the Netherlands. For the C2C return program, we illustrate the performance of the BM solution compared to those of SO and RL.

7.1 Data

We analyze data from our partner retailer from May 2017 to May 2019, consisting of 2.6 million data points. We consider 3 items (items A, B, and C) sold on the webshop of the retailer during a selling season of $T = 60$ days. Customers are allowed to return items to the retailer within $T^{\mathrm{R}} = 30$ days after delivery. Historical data shows that demand is non-stationary and return probabilities depend on the time elapsed since delivery. Demand rate λ_t in day $t = 1, 2, \ldots, T$ and return probability u_i^{R} for an item purchased $i = 1, 2, \ldots, 18$ days ago are as reported in Tables 1 and 2. Item A has low demand (expected demand $\lambda = 1.03$ unit per day) and high returns ($p^{\mathrm{R}} = 0.42$), item B has high demand ($\lambda = 3.77$) and high returns ($p^{\mathrm{R}} = 0.44$); and item C has high demand ($\lambda = 3.68$) and low returns ($p^{\mathrm{R}} = 0.28$). For items A, B, and C, return probabilities u_i^{R} are negligible for $i = 18, \ldots, T^{\mathrm{R}}$. Based on commercial prices, we set shipment and handling costs as $S^{\mathrm{W2C}} =$ €6 and $S^{\mathrm{C2W}} =$ €8. Items A, B, and C are sold for €34.99, €29.99, and €19.99, respectively.

Table 1 Demand rate λ_t for $t = 0, 1, \ldots, T-1$

	Day t (from – to)											
Item	0 – 4	5 – 9	10 – 14	15 – 19	20 – 24	25 – 29	30 – 34	35 – 39	40 – 44	45 – 49	50 – 54	55 – 59
A	1	1	1	0.6	0.4	0.6	0.8	1	1.2	3.4	1	0.4
B	2.4	4.4	5.4	5.6	4	2	5.6	1.6	3.6	4.6	2.8	3.2
C	5	1.8	0.4	1.8	1.4	2	2.2	9.2	2.2	2.4	1.4	14.4

Table 2 Return probability u_i^{R} for $i = 1, 2, \ldots, T^{\mathrm{R}}$

	Day i																	
Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
A	0.06	0.07	0.06	0.06	0.05	0.04	0.04	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0
B	0.06	0.07	0.06	0.06	0.05	0.04	0.04	0.03	0.03	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01
C	0.04	0.04	0.04	0.04	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0	0	0

We note that C2C demand and return characteristics are currently unknown since this concept has never been used before. We consider C2C returns to be fully refunded to encourage customers to join the C2C return program. Based on experts' knowledge, we define optimistic and pessimistic scenarios for the C2C return program as given in Table 3. Solution methods SO and RL can incorporate any functional form for the relation between C2C demand and discount level a. For the sake of simplicity, we use Assumption 2 with $q_0 = 0.05$.

In an optimistic (O) scenario, W2C demand remains the same and total customer demand increases by C2C demand. This corresponds to the best-case demand as described in Section 5.2.3, where λ_t is as given in Table 1 for $t = 0, 1, \ldots, T-1$. The time window for handing-in is set to a reasonable level, i.e., $T^{\mathrm{H}} = 7$ days. As a consequence, delivery lead time for C2C purchases would be reasonable (7 days maximum), paving the way for the best-case demand. In order to encourage returning customers to participate to C2C, we consider a short time window for matching by setting $T^{\mathrm{M}} = 3$ days. Under a short time window for matching and reasonable time window for handing-in, the majority of returning customers (75%) is assumed to perform a C2C return, where we set $u_i^{\mathrm{C2C}} = \gamma \times u_i^{\mathrm{R}}$ and $u_i^{\mathrm{C2W}} = (1-\gamma) \times u_i^{\mathrm{R}}$ with $\gamma = 0.75$ and u_i^{R} is as given in Table 2 for $i = 1, 2, \ldots, T^{\mathrm{R}}$. In an optimistic scenario, a delivery from C2C incurs $S^{\mathrm{C2C}} =$ €4, which is lower than a standard W2C shipment. Return probabilities for W2C and C2C demands are considered as the same, i.e., $u_i^{\mathrm{R}} = u_i^{\mathrm{FR}}$ for $i = 1, 2, \ldots, T^{\mathrm{R}}$.

In a pessimistic (P) scenario, total customer demand remains the same, i.e., W2C demand decreases by C2C demand. We model this situation as the worst-case demand described in Section 5.2.3. In this scenario, the time window for handing-in is set to $T^{\mathrm{H}} = 14$ days, which could cause a long delivery lead time for customers. We impose a long time window for matching by setting $T^{\mathrm{M}} = 5$ days. Due to the hassle a long matching period creates for returning customers, we consider that the majority of returning customers (75%) choose to perform C2W returns, i.e. we set $\gamma = 0.25$. In a pessimistic scenario, a delivery from C2C incurs $S^{\mathrm{C2C}} =$ €6. In addition, return probabilities for C2C demand are set to be much higher than those of W2C demand, where $u_i^{\mathrm{FR}} = 1.5 \times u_i^{\mathrm{R}}$ for $i = 1, 2, \ldots, T^{\mathrm{R}}$.

7.2 Algorithm settings

Solution methods BM, SO, and RL presented in Section 6 have been coded in C++ and experiments are carried out on the Lisa cluster computer installed and maintained by SURFsara, the Netherlands (SURFsara 2021). First, we determine discount level a^{BM} using BM for best-case and worst-case scenarios (see Theorem 1 and 3) and evaluate the performance of the corresponding constant-discount-level policy by simulation. Second, we obtain discount level a^{SO} using SO with $N = 20$. Third, we evaluate non-stationary stochastic policies with RL. For RL, we define the set of possible discount levels as $A = \{0.0, 0.15, 0.25, 0.50\}$.

Table 3 Parameters of optimistic and pessimistic scenarios

Inputs		Scenarios	
		Optimistic (O)	Pessimistic (P)
Demand		Best-case	Worst-case
Returning customers' participation ratio	γ	0.75	0.25
Time window for handing-in (days)	T^{H}	7	14
Time window for matching (days)	T^{M}	3	5
Return rate for C2C purchases (€)	u_i^{FR}	μ^{R}	$1.5 \times \mu^{\mathrm{R}}$
C2C shipment cost (€)	S^{C2C}	4	6

We restrict the per-state computational complexity $(|A|C)^H$ between 160,000 and 250,000 by taking $H=2$, $C=120$; $H=3$, $C=15$; and $H=4$, $C=5$. We call the corresponding algorithms as H2C120, H3C15, and H4C5, respectively. We note that algorithms H2C120, H3C15, and H4C5 are in increasing order of depth and decreasing order of per-state computational complexity. Furthermore, we assess the performance of the system under conventional return program by simulation. For each evaluation, 1,000 replications and common random variables are used. We introduce a cool down period of $T^{\mathrm{R}}=30$ days during which customer returns are allowed but W2C and C2C demands do not occur.

7.3 Results

We define the relative difference in expected total profit of different solutions compared to BM as

$$\Delta\Pi^{\mathrm{SM}} = \frac{\Pi^{\mathrm{SM}} - \Pi^{\mathrm{BM}}}{\Pi^{\mathrm{BM}}}, \tag{7}$$

where Π^{SM} and Π^{BM} are the expected total profits obtained by solution methods SM $\in$ {SO, H2C120, H3C15, H4C5} and BM, respectively. For our numerical experiments, $\Pi^{\mathrm{BM}}>0$ and $\Pi^{\mathrm{SM}}>0$ for all SM $\in$ {SO, H2C120, H3C15, H4C5}. With one-tailed paired t-tests, we check whether we can reject the null hypothesis $\Pi^{\mathrm{BM}}=\Pi^{\mathrm{SM}}$ in favor of $\Pi^{\mathrm{BM}}<\Pi^{\mathrm{SM}}$ (resp. $\Pi^{\mathrm{BM}}>\Pi^{\mathrm{SM}}$) at the significance level of 5% for the cases where $\Delta\Pi^{\mathrm{SM}}>0$ (resp. $\Delta\Pi^{\mathrm{SM}}<0$).

Table 4 shows that BM and SO provide very similar solutions. The relative difference in expected total profit $\Delta\Pi^{\mathrm{SO}}$ is negligible (0.16% at most) even when there is a reasonable difference between constant-discount-levels a^{BM} and a^{SO} ($a^{\mathrm{BM}}-a^{\mathrm{SO}}=3.3\%$ at most for Item A - pessimistic scenario). We observe that the expected total profit is not very sensitive to the discount level.

We observe that SO outperforms the RL algorithms. The trade-off between depth H and sample size C is case-specific. The RL algorithms lead to non-stationary stochastic policies, which are more general than constant-discount-level policies considered in BM and SO. However, the proposed approach is computationally expensive and does not guarantee near-optimality. The computation time required for the evaluation of the RL algorithms is 34 hours on average per instance. SO evaluates the finite set of discount levels within 40 seconds for each instance. Indeed, SO works offline and provides a policy that is easy to implement and understand. The RL method is an online approach, i.e., it needs to be computed at the beginning of each day. The resulting policy might lead to a different discount level for each day, which may also be perceived negatively by the customers.

We measure the value of the C2C return program in terms of expected total profit by

$$\Delta\Pi^{\mathrm{C2C}} = \frac{\Pi^{\mathrm{C2C}} - \Pi^{\mathrm{CON}}}{\Pi^{\mathrm{CON}}}, \tag{8}$$

where Π^{CON} is the expected total profit under the conventional return program and $\Pi^{\mathrm{C2C}} = \max\{\Pi^{\mathrm{BM}}, \Pi^{\mathrm{SO}}, \Pi^{\mathrm{H2C120}}, \Pi^{\mathrm{H3C15}}, \Pi^{\mathrm{H4C5}}\}$. In our numerical experiments, Π^{CON}, we always have $\Pi^{\mathrm{C2C}}>0$.

We define the return rate as the ratio of the number of C2W returns to total demand. The total demand consists of W2C demand (conventional purchase) and C2C demand (C2C purchase). The number of C2W returns are of items delivered from W2C or from C2C. The value of the C2C return program in terms of expected return rate $\Delta\rho^{\text{C2C}}$ is measured by

$$\Delta\rho^{\text{C2C}} = \frac{\rho^{\text{CON}} - \rho^{\text{C2C}}}{\rho^{\text{CON}}}, \tag{9}$$

where ρ^{C2C} and ρ^{CON} are the expected return rates under the conventional and the C2C return programs, respectively. For the C2C return program, we consider the solution with the greatest expected total profit.

Table 4 Performance of solution methods

Item	Setting	a^{BM} (%)	a^{SO} (%)	$\Delta\Pi^{\text{SO}}$ (%)	$\Delta\Pi^{\text{H2C120}}$ (%)	$\Delta\Pi^{\text{H3C15}}$ (%)	$\Delta\Pi^{\text{H4C5}}$ (%)
A	O	31.9	30	0.03	-2.55	-1.23	-2.13
	P	1.7	5	0.16	0.06*	0.01*	-0.05*
B	O	31.2	30	0.01	-2.27	-1.18	-1.91
	P	4.6	5	-0.08	-0.26	-0.26	-0.37
C	O	33.1	35	-0.04	-1.64	-1.03	-1.51
	P	11.6	10	0.07	-0.41	-0.42	-0.40

* H_0 is not rejected.

Table 5 Value of the C2C return program

Item	Setting	Π^{CON} (€/day)	Π^{C2C} (€/day)	$\Delta\Pi^{\text{C2C}}$ (%)	ρ^{CON} (%)	ρ^{C2C} (%)	$\Delta\rho^{\text{C2C}}$ (%)
A	O	7.59	9.48	24.97	41.81	24.13	42.30
	P	7.59	7.61	0.35*	41.81	39.48	5.59
B	O	17.94	23.97	33.61	44.29	25.00	43.56
	P	17.94	18.13	1.08	44.29	39.54	10.72
C	O	15.23	17.89	17.43	27.77	17.12	38.34
	P	15.23	15.55	2.06	27.77	24.74	10.91

* H_0 is not rejected.

As shown in Table 5, in optimistic scenarios, the value of the C2C return program is significant both in terms of expected profit and expected return rate. We observe an increase in expected profit of as much as 34% (from a daily profit of €18 to €24) and a reduction in expected return rate of as much as 44% (the ratio of C2W deliveries to total demand drops from 44% to 25%). We note that the expected profits reported in Table 5 does not include the cost of ordering, purchasing, and inventory holding since we assume them to be the same for both return programs.

In optimistic scenarios, the total demand increases by 12%–24% since C2C always generates additional demand. However, the system can be interpreted as more environmentally friendly considering the ratio of C2W deliveries to total demand. In pessimistic scenarios, the total demand remains the same. In the most pessimistic scenario (Item A), expected total profits Π^{C2C} and Π^{CON} are not significantly different. (One-tailed paired t-tests show that we cannot reject the null hypothesis $\Pi^{\mathrm{C2C}} = \Pi^{\mathrm{CON}}$ in favor of $\Pi^{\mathrm{CON}} < \Pi^{\mathrm{C2C}}$ at the significance level of 5%.) However, the relative reduction in expected return rate is 6%. This shows that the C2C return program can help with reducing return rates, hence provide a more environmentally friendly system, even if the increase in profit is not significant.

8 Numerical Experiments

In this section, we extend our numerical experiments on the performance of the BM and the value of the C2C return program to a broad set of instances to claim generality. We focus on the performance of SO compared to BM due to its good performance, as reported in Section 7.

The setup of the experiments is in accordance with the optimistic and pessimistic scenarios presented in Section 7. Algorithm settings are as given in Section 7.2. We consider the same assumptions and performance measures (see (7), (8), (9)). We fix the selling price $P =$ €20, consider time-independent return probabilities $u_i^{\mathrm{R}} = 0.02$ for $i = 1, 2, \ldots, T^{\mathrm{R}}$ with $T^{\mathrm{R}} = 30$ days ($p^{\mathrm{R}} = 0.45$) and constant demand rates λ. Input parameters that are varied are as reported in Table 6. Performing a full factorial analysis with these input parameters, we obtain 384 instances.

Table 6 reports average (avr.) values over all instances fixing input parameters given in the corresponding rows. By one-tailed paired t-tests, we check whether the null hypothesis H_0: $\Pi^{\mathrm{BM}} = \Pi^{\mathrm{SO}}$ can be rejected in favor of $\Pi^{\mathrm{BM}} > \Pi^{\mathrm{SO}}$ at the significance level of 5%. The 5$^{\text{th}}$ column in Table 6 reports $\Delta\Pi^{\mathrm{SO}}$ for the instances for which H_0 cannot be rejected. The 6$^{\text{th}}$ column in Table 6 reports the number of instances for which the null hypothesis cannot be rejected. We note that, for the instances for which H_0 can be rejected the average $\Delta\Pi^{\mathrm{SO}}$ is -0.12%.

In our numerical experiments, SO and BM provide similar discount levels and expected total profits. For the instances in which SO outperforms BM, the relative difference in expected total profit is on average 0.45% with a maximum of 3.11%. Discount level a^{SO} is very similar to a^{BM} on average (20.81% vs. 20.94%) but the difference between the minimum and the maximum values of a^{SO} is higher than that of a^{BM} (5.00%—40.00% vs. 7.92%—33.51%). Both solutions often behave similarly. The proposed discount level increases with the hand-in time window T^{H} and decreases with the C2C shipment cost S^{C2C} and final return probabilities u_i^{FR}. Indeed, if operational costs for the C2C return program get lower (higher), higher (lower) discounts can be offered. The proposed discount levels decrease with time window for matching T^{M}. This

Table 6 Performance of SO and the value of the C2C return program

Inputs		a^{BM} (avr. %)	a^{SO} (avr. %)	$\Delta\Pi^{\mathrm{SO}}$ (avr. %)*	H_0 not rejected /total instances	$\Delta\Pi^{\mathrm{C2C}}$ (avr. %)	$\Delta\rho^{\mathrm{C2C}}$ (avr. %)
Demand assumption	Best-case	23.61	23.46	0.07	106/192	58.93	24.92
	Worst-case	18.26	18.15	0.66	187/192	36.65	21.51
γ	0.75	20.60	21.69	0.37	140/192	70.16	33.43
	0.25	21.27	19.92	0.52	153/192	25.42	13.00
T	60	20.94	21.12	0.53	153/192	43.61	21.23
	180	20.94	20.49	0.35	140/192	51.97	25.20
T^{H}	14	23.11	22.89	0.54	141/192	55.42	24.28
	7	18.77	18.72	0.36	152/192	40.16	22.15
T^{M}	5	18.78	18.83	0.53	154/192	52.88	24.53
	3	23.09	22.79	0.35	139/192	42.70	21.90
u_i^{FR}	u_i^{R}	23.98	23.39	0.51	164/192	65.38	26.52
	$1.5 \times u_i^{\mathrm{R}}$	17.90	18.23	0.36	129/192	30.20	19.91
S^{C2C}	4	23.48	23.18	0.52	152/192	56.72	24.41
	6	18.40	18.44	0.37	141/192	38.86	22.02
λ	5	20.94	19.10	0.65	99/128	51.37	25.13
	3	20.94	20.23	0.28	97/128	48.89	23.92
	1	20.94	23.09	0.41	97/128	43.11	20.60

* For the instances for which H_0 is not rejected.

is because the likelihood of observing C2C sales gets higher for longer time window for matching. We note that a^{BM} is not affected by the changes in horizon length T or demand rate λ, by definition.

We observe that expected total profits Π^{C2C} and Π^{CON} are significantly different for all instances. As shown in columns 7-8 of Table 6, the C2C return program can be highly valuable both in terms of expected profit and expected return rate. The C2C return program is more valuable when the selling season is long, C2C customers can tolerate long waiting times, the final return probability is not higher than the initial return probability, or demand rates are high. In the most pessimistic scenario where the increase in profit is 3.81% the decrease in expected return rate is 6.1%.

9 Conclusion

Online returns pose a big problem for retailers all over the world. Handling these returns is costly, putting profit under pressure and contributing to CO_2 emissions. In this paper, we introduce the C2C return program where returns skip the retailers' warehouse and are delivered straight to the next customer. Under the C2C return program, when returning an item, customers are asked to keep the returning item for a few more days. During those days the item is promoted on the retailer's website with a discount and the corresponding saving in the CO_2 emission. When the item is sold the returning customer gets a notification to ship the package. Payments and refunds are handled by the online retailer or by an external operator. A provided quick response (QR) label links the returning customer to the new customer. The new customer inspects the item upon reception, scans the product's QR label on the package, and gives a review of the item. The review is added to the profile of the returning customer where it contributes to his/her reputation. The C2C concept is developed by a consortium of contractors, Bearing Point employees, and academics. The next step is to enter the market. For further explanation of the concept, see https://itgoesforward.com/.

Our paper presents the mathematical models behind the C2C concept. The aim is to determine optimal discount levels to offer, maximizing the expected profit of the retailer. First, we propose a customer-based model and show how to determine a constant-discount-level policy. Second, we formulate the real-life problem as an MDP. Due to the curse of dimensionality, determining the optimal policy is computationally intractable. We employ simulation optimization and reinforcement learning algorithms to find reasonably good solutions. We analyze historical real-life demand and return data from a fashion retailer and assess the performance of different solution methods and the value of the C2C return program under different scenarios.

Our numerical experiments show that the base model performs well compared to simulation optimization and reinforcement learning algorithms. In general, the base model outperforms reinforcement learning. The base model and simulation optimization provide similar solutions. We observe that the expected profit of the retailer is not very sensitive to the discount level. Our extended numerical experiments show that the relative difference in expected profit between the base model and simulation optimization is 0.45% on average, with a maximum of 3.11%.

Both our case study and numerical experiments report significant benefits of the C2C return program. In the most optimistic scenario in our case study, we observe an increase of 34% in expected profit and a reduction of 44% in expected return rate. In pessimistic scenarios, in which the increase in expected profit is not significant, the relative reduction in expected return rate can be as much as 6%. Thus, the C2C return program can make the system more environmentally friendly even when it is not highly cost-effective.

This research showed promising initial results for the C2C concept. Future research can extend our work to consolidate multiple items' demand/return, examine customers' response to discount levels, and incorporate inventory control and items' availability in stock. Another research opportunity might be to revisit the C2C concept. For example, C2C sale requests can be collected in advance and fulfilled when the corresponding item is available for a C2C delivery. This would eliminate matching period and could prevent inconveniences for returning customers to keep the item until it has been sold to the next customer.

Acknowledgments

A part of this work was carried out on the Dutch national e-infrastructure with the support of SURF Cooperative. The authors gratefully acknowledge the technical support of Dr. Twan Dollevoet on the use of Lisa cluster computer.

References

Abdulla, H., J.D. Abbey, M. Ketzenberg. 2022. How consumers value retailer's return policy leniency levers: An empirical investigation. *Prod. Oper. Manag.*, 31 (4), 1719-1733.

Abdulla, H., M. Ketzenberg, J.D. Abbey. 2019. Taking stock of consumer returns: A review and classification of the literature. *J. Oper. Manag.*, 65 (6), 560-605.

Adida, E., G. Perakis. 2009. Dynamic pricing and inventory control: Uncertainty and competition. *Oper. Res.*, 58 (2), 289-302.

Akçay, Y., T. Boyacı, D. Zhang. 2013. Selling with money-back guarantees: The impact on prices, quantities, and retail profitability. *Prod. Oper. Manag.*, 22 (4), 777-791.

Borwein, P., T. Erdélyi. 1995. *Polynomials and Polynomial Inequalities,* 1st ed. Springer, New York, NY, USA.

Cohen, M. C., G. Perakis, R. S. Pindyck. 2021. A simple rule for pricing with limited knowledge of demand. *Mgmt. Sci.*, 67 (3), 1329-199.

de Leeuw, S., B. Minguela-Rata, E. Sabet, J. Boter, R. Sigurardottir. 2016. Trade-offs in managing commercial consumer returns for online apparel retail. *Int. J. Oper. Prod. Mgmt.*, 36 (6), 710-731.

Difrancesco, R.M., A. Huchzermeier, D. Schröder. 2018. Optimizing the return window for online fashion retailers with closed-loop refurbishment. *Omega*, 78 205-221.

Diggins, M.A., C. Chen, J. Chen. 2016. A review: Customer returns in fashion retailing. T. M. Choi, ed., *Analytical Modeling Research in Fashion Business*. Springer, Singapore, 31-48.

Forbes. 2019. Costs of return deliveries in the united states in 2017 and 2020 (in billion u.s. dollars). URL `statista.com`. ID 871365.

Griffis, S. E., S. Rao, T. J. Goldsby, T. T. Niranjan. 2012. The customer consequences of returns in online retailing: An empirical analysis. *J. Oper. Manag.*, 30 (4), 282-294.

Hjort, K., B. Lantz. 2016. The impact of returns policies on profitability: A fashion e-commerce case. *J. Bus. Res.*, 69 (11), 4980-4985.

Hsieh, M. 2021. Customer preferences for customer-to-customer return logistics and the differences in customer preferences between the Netherlands and Taiwan. Master's thesis, Vrije Universiteit Amsterdam, Faculty of Economics and Business Administration. URL `https://www.ubvu.vu.nl/pub/fulltext/scripties/27_2690479_0.pdf`.

Kearns, M., Y. Mansour, Y. N. G. Andrew. 2002. A sparse sampling algorithm for near-optimal planning in large Markov decision processes. *Mach. Learn.*, 49 (2), 193-208.

Lee, D.H. 2015. An alternative explanation of consumer product returns from the postpurchase dissonance and ecological marketing perspectives. *Psychol. Mark.*, 32 (1), 49-64.

Minnema, A., T.H.A. Bijmolt, J.A. Petersen, J.D. Shulman. 2018. Managing product returns within the customer value framework. Robert W. Palmatier, V. Kumar, Colleen M. Harmeling, eds., *Customer Engagement Marketing*. Springer International Publishing, Cham, 95-118. URL `https://doi.org/10.1007/`

978-3-319-61985-9_5.

Mollenkopf, D. A., E. Rabinovich, T. M. Laseter, K. K. Boyer. 2007. Managing internet product returns: A focus on effective service operations. *Decis. Sci.*, 38 (2), 215-250.

Mostard, J., R. Teunter. 2006. The newsboy problem with resalable returns: A single period model and case study. *Eur. J. Oper. Res.*, 169 (1), 81-96.

Narang, U., V. Shankar. 2019. Mobile app introduction and online and offline purchases and product returns. *Mark. Sci.*, 38 (5), 756-772.

Neumark, S. 2014. *Solution of Cubic and Quartic Equations,* 1st ed. Pergamon Press Ltd., Oxford.

NOS. 2020. Retour roemoer. URL https://app.nos.nl/op3/pakjes/index.html. Retrieved August 31, 2020.

Patel, P.C., C. Baldauf, S. Karlsson, P. Oghazi. 2021. The impact of free returns on online purchase behavior: Evidence from an intervention at an online retailer. *J. Oper. Manag.*, (In-press).

Péret, L., F. Garcia. 2004. On-line search for solving markov decision processes via heuristic sampling. *Proceedings of the 16th European Conference on Artificial Intelligence,* IOS press, Amsterdam, The Netherlands. 530-534.

Shang, G., M.E. Ferguson, M.R. Galbreth. 2019. Where should I focus my return reduction efforts? Empirical guidance for retailers. *Decis. Sci.*, 50 (4), 877-909.

Shang, G., P. Pekgün, M. Ferguson, M. Galbreth. 2017. How much do online consumers really value free product returns? Evidence from e-Bay. *J. Oper. Manag.*, 53 45-62.

Shulman, J.D., A.T. Coughlan, R.C. Savaskan. 2009. Optimal restocking fees and information provision in an integrated demand-supply model of product returns. *Manuf. Serv. Oper. Manag*, 11 (4), 577-594.

Statista. 2018. Return rate of online ordered clothing by U.S. consumers 2018. URL statista.com. ID 962119.

SURFsara. 2021. Lisa system. URL https://userinfo.surfsara.nl/systems/lisa/description. Retrieved October 7, 2021.

Vetterling, W.T., W.H. Press. 1992. *Numerical recipes in FORTRAN: The art of scientific computing,* 2nd ed. Cambridge University Press, New York, NY, USA.

Wiersma, S. 2021. Customer-to-customer return logistics: exploring the consumer perspective to a new approach for managing returns. Master's thesis, Vrije Universiteit Amsterdam, Faculty of Economics and Business Administration. URL https://www.ubvu.vu.nl/pub/fulltext/scripties/27_2693468_0.pdf.

Xing, W., L. Dan, C. Jing, L. Yongqin. 2020. Managing customer returns strategy with the option of selling returned products. *Int. J. Prod. Econ.*, 230 107794.

Appendix: Proofs and Theorems

Proof of Theorem 1 Under Assumption 1, substituting $p^{\mathrm{C2W}} = p^{\mathrm{R}} - p^{\mathrm{C2C}}$ into equation (1) and writing the difference between equation (2) and equation (1), we have:

$$E\left[\Pi_{\mathrm{B}}(a)\right] - E\left[\Pi\right] = p^{\mathrm{C2C}} p^{\mathrm{M}}(a) \Big(P\left(1-a\right)\left(1-p^{\mathrm{FR}}\right) - \phi_B \Big). \tag{EC.1}$$

Note that $p^{\mathrm{C2C}} p^{\mathrm{M}}(a)$ is non-negative for all $a \in [0,1]$. Hence it follows that $E\left[\Pi_{\mathrm{B}}(a)\right] \geq E\left[\Pi\right]$ if and only if

$$P\left(1-a\right)\left(1-p^{\mathrm{FR}}\right) \geq \phi_B.$$

Similarly, by taking the difference of $E\left[\Pi_{\mathrm{W}}(a)\right]$ and $E\left[\Pi\right]$, we get:

$$E\left[\Pi_{\mathrm{W}}(a)\right] - E\left[\Pi\right] = p^{\mathrm{C2C}} p^{\mathrm{M}}(a) \Big(P\left(1-a\right)\left(1-p^{\mathrm{FR}}\right) - \phi_B - E\left[\Pi\right] \Big) = p^{\mathrm{C2C}} p^{\mathrm{M}}(a) \Big(P\left(1-a\right)\left(1-p^{\mathrm{FR}}\right) - \phi_W \Big). \tag{EC.2}$$

Again, note that $p^{\mathrm{C2C}} p^{\mathrm{M}}(a)$ is non-negative for all $a \in [0,1]$, and hence we find that $E\left[\Pi_{\mathrm{W}}(a)\right] \geq E\left[\Pi\right]$ if and only if

$$P\left(1-a\right)\left(1-p^{\mathrm{FR}}\right) \geq \phi_W.$$

□

The parameter

$$\beta = \frac{\left(T^{\mathrm{M}}-1\right)\phi_B}{\left(T^{\mathrm{M}}+1\right)\left(1-p^{\mathrm{FR}}\right)P}. \tag{EC.3}$$

plays an important role in characterizing the shape (convex or concave) of the profit functions and hence in optimizing the discount level. The next lemma will be used in the proofs of Theorem EC.1 and Theorem 2.

LEMMA EC.1. *Under Assumption 2 and the best-case scenario, the expected profit function of the C2C return program $\mathbb{E}\left[\Pi_B(a)\right]$ can be characterized as follows:*

i. If $\beta \leq 0$, $\mathbb{E}\left[\Pi_B(a)\right]$ is concave on $a \in [0,1]$,

ii. If $0 < \beta < 1$, $\mathbb{E}\left[\Pi_B(a)\right]$ is concave on $a \in [0, 1-\beta]$ and convex on $a \in [1-\beta, 1]$,

iii. If $\beta \geq 1$, $\mathbb{E}\left[\Pi_B(a)\right]$ is convex on $a \in [0,1]$.

Let $T^M \in \mathbb{Z}^+ \setminus \{1\}$. Under Assumption 2 and the worst-case scenario, the expected profit function of the C2C return program $\mathbb{E}\left[\Pi_W(a)\right]$ can be characterized as follows:

i. If $\phi_W \leq 0$, $\mathbb{E}\left[\Pi_W(a)\right]$ is concave on $a \in [0,1]$,

ii. if $\phi_W \geq 3P\left(1-p^{FR}\right)$, $\mathbb{E}\left[\Pi_W(a)\right]$ is convex on $a \in [0,1]$,

iii. if $0 < \phi_W < 3P\left(1-p^{FR}\right)$, the second-order condition is inconclusive to identify whether $\mathbb{E}\left[\Pi_W(a)\right]$ is convex or concave on $a \in [0,1]$.

Proof of Lemma EC.1: Under Assumption 2 the second order derivative of $E\left[\Pi_{\mathrm{B}}(a)\right]$ equals

$$\frac{d^2E\left[\Pi_{\mathrm{B}}(a)\right]}{da^2}=p^{\mathrm{C2C}}\left(1-a\right)^{T^{\mathrm{M}}-2}\left(1-q_0\right)^{T^{\mathrm{M}}}T^{\mathrm{M}}\bigg(\left(T^{\mathrm{M}}-1\right)\phi_B-\left(1-a\right)\left(T^{\mathrm{M}}+1\right)\left(1-p^{\mathrm{FR}}\right)P\bigg),$$

where $\phi_B=C^{\mathrm{HI}}p^{\mathrm{HI}}+C^{\mathrm{LHI}}\left(1-p^{\mathrm{HI}}\right)-\left(1-p^{\mathrm{FR}}\right)S^{C2W}$.

Since $p^{\mathrm{C2C}}\left(1-q_0\right)^{T^{\mathrm{M}}}T^{\mathrm{M}}$ is a positive constant, $E\left[\Pi_{\mathrm{B}}(a)\right]$ is concave (resp. convex) on an interval of $a\in[0,1]$ if and only if

$$\left(1-a\right)^{T^{\mathrm{M}}-2}\bigg(\left(T^{\mathrm{M}}-1\right)\phi_B-\left(1-a\right)\left(T^{\mathrm{M}}+1\right)\left(1-p^{\mathrm{FR}}\right)P\bigg) \qquad \text{(EC.4)}$$

is non-positive (resp. non-negative) on this interval. Since $0\le p^{\mathrm{FR}}<1$, the sign of (EC.4) is the same as the sign of

$$\left(1-a\right)^{T^{\mathrm{M}}-2}\left[\beta-\left(1-a\right)\right], \qquad \text{(EC.5)}$$

where

$$\beta=\frac{\left(T^{\mathrm{M}}-1\right)\phi_B}{\left(T^{\mathrm{M}}+1\right)\left(1-p^{\mathrm{FR}}\right)P}.$$

As $\left(1-a\right)^{T^{\mathrm{M}}-2}\ge 0$ for $a\in[0,1]$, the sign of (EC.5) is defined by the sign of $[\beta-(1-a)]$. So, we consider the following cases to characterize the sign of the corresponding expression:

i. If $\beta\le 0$, then (EC.5) is non-positive. Thus, $E\left[\Pi_{\mathrm{B}}(a)\right]$ is concave on $a\in[0,1]$,

ii. If $0<\beta<1$, then (EC.5) non-positive for $a\in[0,1-\beta]$ and (EC.5) is non-negative for $a\in[1-\beta,1]$. Thus, $E\left[\Pi_{\mathrm{B}}(a)\right]$ is concave on $a\in[0,1-\beta]$ and convex on $a\in[1-\beta,1]$,

iii. If $\beta\ge 1$, then (EC.5) is non-negative. Thus, $E\left[\Pi_{\mathrm{B}}(a)\right]$ is convex on $a\in[0,1]$.

Under Assumption 2, the second order derivative of $E\left[\Pi_{\mathrm{W}}(a)\right]$ is

$$\frac{d^2E\left[\Pi_{\mathrm{W}}(a)\right]}{da^2}=\frac{p^{\mathrm{C2C}}(1-q_0)^{T^{\mathrm{M}}}(1-a)^{T^{\mathrm{M}}-2}}{\left(1+p^{\mathrm{C2C}}p^{\mathrm{M}}(a)\right)^3}\bigg[\left(\left(1+p^{\mathrm{C2C}}p^{\mathrm{M}}(a)\right)-\left(1+p^{\mathrm{C2C}}\left(2-p^{\mathrm{M}}(a)\right)\right)T^{\mathrm{M}}\right)\left(P(1-a)\left(1-p^{\mathrm{FR}}\right)-\phi_W\right)$$
$$-2\left(1+p^{\mathrm{C2C}}p^{\mathrm{M}}(a)\right)P(1-a)\left(1-p^{\mathrm{FR}}\right)\bigg].$$

The fraction given in the above expression is non-negative on $a\in[0,1]$. Accordingly, $E\left[\Pi_{\mathrm{W}}(a)\right]$ is concave (resp. convex) on $[0,1]$ if and only if

$$\left(\left(1+p^{\mathrm{C2C}}p^{\mathrm{M}}(a)\right)-\left(1+p^{\mathrm{C2C}}\left(2-p^{\mathrm{M}}(a)\right)\right)T^{\mathrm{M}}\right)\left(P(1-a)\left(1-p^{\mathrm{FR}}\right)-\phi_W\right)-2\left(1+p^{\mathrm{C2C}}p^{\mathrm{M}}(a)\right)P(1-a)\left(1-p^{\mathrm{FR}}\right) \qquad \text{(EC.6)}$$

is non-positive (resp. non-negative) for any given $a\in[0,1]$. For any given a where $a\in[0,1)$, the sign of (EC.6) is the same as the sign of

$$\left(1-\frac{\left(1+p^{\mathrm{C2C}}\left(2-p^{\mathrm{M}}(a)\right)\right)T^{\mathrm{M}}}{\left(1+p^{\mathrm{C2C}}\right)p^{\mathrm{M}}(a)}\right)\left(\frac{P(1-a)\left(1-p^{\mathrm{FR}}\right)-\phi_W}{2P(1-a)\left(1-p^{\mathrm{FR}}\right)}\right)-1. \qquad \text{(EC.7)}$$

Under the assumption that $T^{\mathrm{M}} \in \mathbb{Z}^+ \setminus \{1\}$, the first part of (EC.7) is at most -1. Correspondingly, if

$$\left(\frac{P(1-a)\left(1-p^{\mathrm{FR}}\right)-\phi_W}{2P(1-a)\left(1-p^{\mathrm{FR}}\right)}\right) \tag{EC.8}$$

is greater (smaller) than 0 (-1), the corresponding profit function is concave (convex).

Considering the condition given in (EC.8) for $a \in [0,1)$, we can characterize the corresponding profit function as follows:

i. If $\phi_W \leq 0$, then the expression given in (EC.8) has a positive sign and so equation (EC.7) is non-positive. Thus, we can conclude that $E\left[\Pi_{\mathrm{B}}(a)\right]$ is concave on $a \in [0,1]$.

ii. If $\phi_W \geq 3P(1-p^{FR})$, then the expression given in (EC.8) is at most -1. Thus, we can conclude that equation (EC.7) is non-negative and hence $E\left[\Pi_{\mathrm{B}}(a)\right]$ is convex on $a \in [0,1]$.

iii. If $0 < \phi_W < 3P(1-p^{FR})$, then the expression given in equation (EC.8) is not sufficient to analytically determine the sign of the second derivative, and hence to identify whether the corresponding function is concave or convex.

□

THEOREM EC.1. *Under Assumption 2 and the best-case scenario, the optimal discount level a_B^* maximizing the expected profit function of the C2C return program $\mathbb{E}\left[\Pi_B(a)\right]$ is obtained as follows:*

i. If $\beta < 1$ and there exists an $\hat{a} \in [0,1]$ satisfying

$$T^M \phi_B \left(1-\hat{a}\right)^{T^M-1} - \left(T^M+1\right)\left(1-p^{FR}\right) P \left(1-\hat{a}\right)^{T^M} + \frac{\left(1-p^{FR}\right) P}{\left(1-q_0\right)^{T^M}} = 0, \tag{EC.9}$$

then the optimal discount level is a_B^*=$\hat{a}$,

ii. Otherwise, the optimal discount level is $a_B^* = 0$.

Let $T^M \in \mathbb{Z}^+ \setminus \{1\}$. Under Assumption 2 and the worst-case scenario, the optimal discount level a_W^* maximizing the expected profit function of the C2C return program $\mathbb{E}\left[\Pi^W(a)\right]$ is obtained as follows:

i. If $\phi_W \leq 0$ and there exists an $\tilde{a} \in [0,1]$ satisfying

$$\frac{\left(1-p^M(\tilde{a})\right) T^M \left(\phi_W + P\left(1-\tilde{a}\right)\left(1-p^{FR}\right)\right)}{\left(1+p^{C2C} p^M(\tilde{a})\right)^2} - \frac{P\left(1-p^{FR}\right) p^M(\tilde{a})\left(1-\tilde{a}\right)}{1+p^{C2C} p^M(\tilde{a})} = 0, \tag{EC.10}$$

then the optimal discount level is a_W^*=$\tilde{a}$.

ii. If $\phi_W \leq 0$ and $\tilde{a} \in [0,1]$ satisfying (EC.10) does not exist, or if $\phi_W \geq 3P\left(1-p^{FR}\right)$, then the optimal discount level is $a_W^* = 0$.

iii. If $0 < \phi_W < 3P\left(1-p^{FR}\right)$, then the first- and second-order conditions are not sufficient for characterizing the optimal discount level a_W^* analytically.

Proof of Theorem EC.1: Under Assumption 2, we can derive the optimal discount level a_{B}^* by using the cases introduced for $E\left[\Pi_{\mathrm{B}}(a)\right]$ in Lemma EC.1.

i. Let $\beta \le 0$. In this case, $E[\Pi_{\mathrm{B}}(a)]$ is concave on $a \in [0,1]$ from Lemma EC.1. The first derivative of the expected profit function $E[\Pi_{\mathrm{B}}(a)]$ is:

$$\frac{dE[\Pi_{\mathrm{B}}(a)]}{da} = -p^{\mathrm{C2C}}\Big((1-a)^{T^{\mathrm{M}}-1}(1-q_0)^{T^{\mathrm{M}}}\big[T^{\mathrm{M}}\phi_B - \big(T^{\mathrm{M}}+1\big)\big(1-p^{\mathrm{FR}}\big)P(1-a)\big] + \big(1-p^{\mathrm{FR}}\big)P\Big). \tag{EC.11}$$

Note that p^{C2C} is a non-negative constant and $0 \le q_0 < 1$. From the first-order condition, if there exists $\hat{a} \in [0,1]$ satisfying,

$$T^{\mathrm{M}}\phi_B(1-\hat{a})^{T^{\mathrm{M}}-1} - \big(T^{\mathrm{M}}+1\big)\big(1-p^{\mathrm{FR}}\big)P(1-\hat{a})^{T^{\mathrm{M}}} + \frac{(1-p^{\mathrm{FR}})P}{(1-q_0)^{T^{\mathrm{M}}}} = 0, \tag{EC.12}$$

then $a^*_{\mathrm{B}} = \hat{a}$. Note that (EC.11) is non-positive for $a = 1$. Therefore, if $\hat{a} \in [0,1]$ satisfying (EC.12) does not exist, $E[\Pi(a)]$ is a decreasing concave function on $a \in [0,1]$. Thus, $a^*_{\mathrm{B}} = 0$.

ii. Let $0 < \beta < 1$. In this case, $\mathbb{E}[\Pi_{\mathrm{B}}(a)]$ is concave on $a \in [0, 1-\beta]$ and convex on $a \in [1-\beta, 1]$. To determine a^*_{B}, we should consider the first-order condition on the concave part $a \in [0, 1-\beta]$ and the boundaries of $1-\beta$ and 1 on the convex part $a \in [1-\beta, 1]$. If there exist $\hat{a} \in [0, 1-\beta]$ satisfying (EC.12), then

$$a^*_{\mathrm{B}} = \arg\max\{E[\Pi_{\mathrm{B}}(\hat{a})], E[\Pi_{\mathrm{B}}(1)]\}.$$

For $a \in [0,1]$, $\beta > 0$ implies $\phi_B > 0$. So, we get

$$E[\Pi_{\mathrm{B}}(a)] - E[\Pi(1)] = p^{\mathrm{M}}(a)\big(1-p^{\mathrm{FR}}\big)P(1-a) + \phi_B\big(1-p^{\mathrm{M}}(a)\big) \ge 0, \tag{EC.13}$$

which implies that at $a = 1$ the minimal profit is obtained. Thus, we can conclude that $a^*_{\mathrm{B}} = \hat{a}$.

One can verify that (EC.11) is non-positive for $a = 1-\beta$. Therefore, if $\hat{a} \in [0, 1-\beta]$ satisfying (EC.12) does not exist, $E[\Pi_{\mathrm{B}}(a)]$ is a decreasing concave function on $a \in [0, 1-\beta]$. From (EC.13), $a = 1$ cannot be optimal. Thus, we can conclude that $a^*_{\mathrm{B}} = 0$.

iii. Let $\beta \ge 1$. In this case, $E[\Pi_{\mathrm{B}}(a)]$ is convex on $a \in [0,1]$ and we should consider the boundaries 0 and 1. From (EC.13), $a^*_{\mathrm{B}} = 0$.

Finally, by merging cases (i)–(iii) of this proof, Theorem EC.1 (i)–(ii) follows.

Similarly, using the cases introduced for $E[\Pi_{\mathrm{W}}(a)]$ in Lemma EC.1, we can characterize the optimal discount level a^*_{W} as follows:

i. Let $\phi_W \le 0$. In this case, $E[\Pi_{\mathrm{W}}(a)]$ is concave on $a \in [0,1]$ from Lemma EC.1. The first derivative of the corresponding profit function is:

$$\frac{dE[\Pi_{\mathrm{W}}(a)]}{da} = p^{\mathrm{C2C}}\left(\frac{(1-p^{\mathrm{M}}(a))T^{\mathrm{M}}(P(1-a)(1-p^{\mathrm{FR}}) - \phi_W)}{(1-a)(1+p^{\mathrm{C2C}}p^{\mathrm{M}}(a))^2} - \frac{P(1-p^{\mathrm{FR}})p^{\mathrm{M}}(a)}{(1+p^{\mathrm{C2C}}p^{\mathrm{M}}(a))}\right). \tag{EC.14}$$

Note that p^{C2C} is a positive constant. From the first-order condition, if there exists $\tilde{a} \in [0,1]$ satisfying,

$$\frac{(1-p^{\mathrm{M}}(\tilde{a}))T^{\mathrm{M}}(P(1-\tilde{a})(1-p^{\mathrm{FR}}) - \phi_W)}{(1-\tilde{a})(1+p^{\mathrm{C2C}}p^{\mathrm{M}}(\tilde{a}))^2} = \frac{P(1-p^{\mathrm{FR}})p^{\mathrm{M}}(\tilde{a})}{(1+p^{\mathrm{C2C}}p^{\mathrm{M}}(\tilde{a}))}. \tag{EC.15}$$

then $a^*_{\mathrm{W}} = \tilde{a}$.

By using the fact that $p^{\mathrm{M}}(a) = 1-(1-q(a))^{T^{\mathrm{M}}}$ and $q(a) = q_0 + a(1-q_0)$ (according to Assumption 2), it can be verified that (EC.14) is non-positive for $a = 1$. Therefore, if $\tilde{a} \in [0,1]$ satisfying (EC.15) does not exist, $E[\Pi_{\mathrm{W}}(a)]$ is a decreasing concave function on $a \in [0,1]$. Thus, $a^*_{\mathrm{W}} = 0$.

ii. Let $\phi_W \geq 3P(1-p^{FR})$. In this case, $E[\Pi_{\mathrm{W}}(a)]$ is convex on $a \in [0,1]$ from Lemma EC.1. We should consider the boundaries of 0 and 1 to find the optimal discount level. Correspondingly, we have:

$$a^*_{\mathrm{W}} = \arg\max\left\{E[\Pi_{\mathrm{W}}(0)], E[\Pi_{\mathrm{W}}(1)]\right\}. \tag{EC.16}$$

By taking the difference between $E[\Pi_{\mathrm{W}}(0)]$ and $E[\Pi_{\mathrm{W}}(1)]$, it is possible to determine the conditions that identify which of these two discount levels would be the solution to the optimization problem given in (EC.16). It follows that:

$$E[\Pi_{\mathrm{W}}(0)] - E[\Pi_{\mathrm{W}}(1)] = \frac{-p^{\mathrm{C2C}}P(1+p^{\mathrm{C2C}})(1-p^{\mathrm{FR}}) + p^{\mathrm{C2C}}(1-q_0)^{T^{\mathrm{M}}}\left(P(1-p^{\mathrm{FR}})(1+p^{\mathrm{C2C}}) - \phi_W\right)}{-(1+p^{\mathrm{C2C}})(1+p^{\mathrm{C2C}}(1-(1-q_0)^{T^{\mathrm{M}}}))}. \tag{EC.17}$$

As $\phi_W \geq 3P(1-p^{FR})$, the expression given in (EC.17) has a positive sign. This implies that $a = 0$ is more profitable than $a = 1$, and we can conclude that $a^*_{\mathrm{W}} = 0$.

iii. Finally, let $0 < \phi_W < 3P(1-p^{FR})$. In this case, we are enable to determine the shape of the function analytically, and hence the first- and second-order conditions are not sufficient for analytically characterizing the optimal discount level a^*_{W}.

□

It follows that under the best-case scenario, one needs to solve the polynomial equation (EC.9). If the degree of the polynomial is at most 4, i.e., if matching time window $T^{\mathrm{M}} \leq 4$, it is well-known that there exists a closed-form solution for finding $\hat{a}$ (see e.g. Vetterling and Press (1992), Borwein and Erdélyi (1995), Neumark (2014), and the references in there). For higher degrees, the optimal discount level can still be found numerically as the function turns out to be unimodal (see Lemma EC.1).

To interpret case ii. under the best-case scenario, note that $\beta > 1$ implies $\phi_B > P(1-p^{\mathrm{FR}})$ (by using equation (EC.3)), meaning that the C2C return program is always worse then the conventional one (see Theorem 1). Hence, we like to have as few C2C customers as possible, which is achieved by setting $a = 0$. Unfortunately, for the worst-case scenario the interpretation of the ranges is less clear and there is a range of ϕ_W for which we cannot determine an analytical expression for the optimal discount level.

Proof of Theorem 2 As follows from the proof of Theorem EC.1, the function $E[\Pi_{\mathrm{B}}(a)]$ is decreasing at $a = 1$. Furthermore, we know from Lemma EC.1 that (in the most general case) the first part of the function $E[\Pi_{\mathrm{B}}(a)]$ is concave and the second part convex. Combining these observations gives the result. □

CONJECTURE EC.1. *The function $E[\Pi_W(a)]$ is unimodal on the interval $[0,1]$.*

There are several reasons why we have this conjecture. First, note that $\phi_W \leq P(1-a)(1-p^{\mathrm{FR}})$ implies that the term in equation (EC.8) is non-negative, which in turn implies that equation (EC.7) is non-positive. Secondly, for $\phi_W \geq 3P(1-a)(1-p^{\mathrm{FR}})$ the term in equation (EC.8) is at most -1, which in turn implies that equation (EC.7) is non-negative. This implies that for any value of ϕ_W the function $E[\Pi_{\mathrm{W}}(a)]$ starts as concave or ends as convex (or both) on the interval $[0,1]$. Moreover, from the proof of Theorem EC.1, the function $E\left[\Pi_{\mathrm{W}}(a)\right]$ is decreasing at $a=1$. Hence, if the function only switches once from concave to convex, then $E[\Pi_{\mathrm{W}}(a)]$ is unimodal. Finally, in all the parameter settings of our experiments, the best solution was always obtained by the golden section search, also suggesting the unimodality of the function.

Proof of Theorem 3 Recall that $\mathbb{E}\left[\Pi_{\mathrm{W}}(a)\right]=\frac{\mathbb{E}[\Pi_{\mathrm{B}}(a)]}{1+p^{\mathrm{C2C}}p^{\mathrm{M}}(a)}$, for which we know from the first part that $E[\Pi_{\mathrm{B}}(a)]$ is unimodal on $[0,1]$. That is, $E[\Pi_{\mathrm{B}}(a)]$ is increasing on $[0,a_B^*]$ and decreasing on $[a_B^*,1]$. Now let us focus on $p^{\mathrm{M}}(a)$ which can be written as

$$p^{\mathrm{M}}(a)=1-(1-q_0)^{T^M}(1-a)^{T^M}.$$

By analysing the first derivative, it turns out that $p^{\mathrm{M}}(a)$ is a positive and increasing function on $[0,1]$, as well as $1+p^{\mathrm{C2C}}p^{\mathrm{M}}(a)$. Since we divide $\Pi_{\mathrm{B}}(a)$ by a positive and increasing function, $\Pi_{\mathrm{W}}(a)$ will be decreasing on $[a_B^*,1]$, which implies that the maximizer a_B^* of $\Pi_{\mathrm{W}}(a)$ should be found in the interval $[0,a_B^*]$, proving the result. □

Proof of Theorem 4 In this proof, we skip time index t due to stationarity under Assumption 3. The long-run average profit of the conventional multi-customer system can be written as

$$\mathbb{E}\left[\hat{\Pi}\right]=R^{W2C}\mathbb{E}\left[d^{\mathrm{W2C}}\right]-C^{C2W}\sum_{i=1}^{T^{\mathrm{R}}}\mathbb{E}\left[r_i^{\mathrm{C2W}}\right]=R^{W2C}\lambda-C^{C2W}\sum_{i=1}^{T^{\mathrm{R}}}\mathbb{E}\left[r_i^{\mathrm{C2W}}\right]. \tag{EC.18}$$

By definition, x_1 is the number of customers recently entered the system, i.e., d^{W2C}. This implies that x_1 follows the same distribution as d^{W2C}, i.e., $x_1 \sim \mathit{Poisson}\,(\lambda)$.

The number of returns r_i^{C2W} follows a binomial distribution with parameters x_i and u_i^{R} for $i \in \{1,2,\ldots,T^{\mathrm{R}}\}$. This allows us to use the splitting property of the Poisson process to characterize the distribution of the returns stemming from x_1, i.e., $r_1^{\mathrm{C2W}} \sim \mathit{Poisson}\,(\lambda u_1^{\mathrm{R}})$. Non-returned items will spend one more day in the system thereby constituting x_2. Similarly, the number of non-returned items follows a Poisson distribution, i.e, $x_2 \sim \mathit{Poisson}\,(\lambda\,(1-u_1^{\mathrm{R}}))$. Advancing this argumentation for more periods, we can characterize the distribution and expectation of r_i^{C2W} as follows

$$r_i^{\mathrm{C2W}} \sim \mathit{Poisson}\left(\lambda u_i^{\mathrm{R}}\prod_{j=1}^{i-1}\left(1-u_j^{\mathrm{R}}\right)\right) \text{ and } \mathbb{E}\left[r_i^{\mathrm{C2W}}\right]=\lambda u_i^{\mathrm{R}}\prod_{j=1}^{i-1}\left(1-u_j^{\mathrm{R}}\right).$$

Incorporating this into (EC.18), we have

$$\mathbb{E}\left[\hat{\Pi}\right]=R^{W2C}\lambda-C^{C2W}\sum_{i=1}^{T^{\mathrm{R}}}\mathbb{E}\left[r_i^{\mathrm{C2W}}\right]=R^{W2C}\lambda-C^{C2W}\lambda\sum_{i=1}^{T^{\mathrm{R}}}u_i^{\mathrm{R}}\prod_{j=1}^{i-1}\left(1-u_j^{\mathrm{R}}\right).$$

This implies

$$\mathbb{E}\left[\hat{\Pi}\right]=\lambda\left(R^{W2C}-C^{C2W}p^{\mathrm{R}}\right)=\lambda\mathbb{E}\left[\Pi\right].$$

□

Proof of Theorem 5 In this proof, we skip time index t due to stationary under Assumptions 3. Considering $T^{\mathrm{M}}=1$, the long-run average profit of the C2C multi-customer system can be written as

$$\begin{aligned}\mathbb{E}\left[\hat{\Pi}\left(a\right)\right]=&R^{\mathrm{W2C}}\mathbb{E}\left[d^{\mathrm{W2C}}\right]-C^{\mathrm{C2W}}\sum_{i=1}^{T^{\mathrm{R}}}\mathbb{E}\left[r_i^{\mathrm{C2W}}\right]-C^{\mathrm{C2W}}\mathbb{E}\left[\left(y_1-d^{\mathrm{C2C}}\right)^+\right]\\&+\left(R^{\mathrm{C2C}}(a)-\left(C^{\mathrm{HI}}p^{\mathrm{HI}}+C^{\mathrm{LHI}}(1-p^{\mathrm{HI}})\right)-C^{\mathrm{C2C}}(a)p^{\mathrm{FR}}\right)\mathbb{E}\left[d^{\mathrm{C2C}}\right].\end{aligned}\tag{EC.19}$$

Under Assumption 3, we have

$$\mathbb{E}\left[d^{\mathrm{W2C}}\right]=\lambda.\tag{EC.20}$$

Following the same reasoning as in the proof of Theorem 4, $x_1\sim Poisson\left(\lambda\right)$. The number of returns (r_i^{C2W}, r_i^{C2C}) follows a multinomial distribution with parameters (x_i, u_i^{C2W}, u_i^{C2C}) for $i\in\{1,2,\ldots,T^{\mathrm{R}}\}$. Thus, we can use the splitting property of the Poisson process to characterize the distributions of the number of C2W and C2C returns stemming from x_1, i.e., $r_1^{\mathrm{C2W}}\sim Poisson\left(\lambda u_1^{\mathrm{C2W}}\right)$ and $r_1^{\mathrm{C2C}}\sim Poisson\left(\lambda u_1^{\mathrm{C2C}}\right)$. The non-returned items will spend one more day in the system and constitute x_2. Similarly, the number of non-returned items follows a Poisson distribution, i.e, $x_2\sim Poisson\left(\lambda\left(1-u_1^{\mathrm{C2C}}-u_1^{\mathrm{C2W}}\right)\right)$. Advancing this argumentation for more periods, we can characterize the distributions and expectations of r_i^{C2W} and r_i^{C2C} as follows

$$r_i^{\mathrm{C2W}}\sim Poisson\left(\lambda u_i^{C2W}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right)\right)\text{ and }\mathbb{E}\left[r_i^{\mathrm{C2W}}\right]=\lambda u_i^{C2W}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right),\tag{EC.21}$$

$$r_i^{\mathrm{C2C}}\sim Poisson\left(\lambda u_i^{C2C}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right)\right)\text{ and }\mathbb{E}\left[r_i^{\mathrm{C2C}}\right]=\lambda u_i^{C2C}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right).\tag{EC.22}$$

Recall that the sum of r_i^{C2C} will constitute y_1. For each $i\in\{1,2,\ldots,T^{\mathrm{R}}\}$, r_i^{C2C} follows a Poisson distribution. Moreover, the number of returns in different periods are independent from each other. Thus, we can use the merging property of the Poisson process to characterize the distribution and expectation of y_1. This leads to

$$y_1\sim Poisson\left(\lambda\sum_{i=1}^{T^{\mathrm{R}}}u_i^{C2C}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right)\right)\text{ and }\mathbb{E}\left[y_1\right]=\lambda\sum_{i=1}^{T^{\mathrm{R}}}u_i^{C2C}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right).\tag{EC.23}$$

According to (EC.23), the number of C2C demands d^{C2C} is a Poisson distributed random variable. We can use the splitting property and derive the distribution and expectation of d^{C2C}. This leads to

$$d^{\mathrm{C2C}}\sim Poisson\left(\lambda q(a)\sum_{i=1}^{T^{\mathrm{R}}}u_i^{C2C}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right)\right),\tag{EC.24}$$

$$\mathbb{E}\left[d^{\mathrm{C2C}}\right]=\lambda q(a)\sum_{i=1}^{T^{\mathrm{R}}}u_i^{C2C}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right).\tag{EC.25}$$

Note that $(y_1 - d^{\mathrm{C2C}})^+$ denotes the number of items that are not sold within the matching period. By definition, d^{C2C} cannot exceed y_1. Thus, $(y_1 - d^{\mathrm{C2C}})^+ = y_1 - d^{\mathrm{C2C}}$. Again, we can use the splitting property and establish the distribution and expectation of $(y_1 - d^{\mathrm{C2C}})^+$ as follows

$$\left(y_1 - d^{\mathrm{C2C}}\right)^+ \sim Poisson\left(\lambda\big(1-q(a)\big)\sum_{i=1}^{T^{\mathrm{R}}} u_i^{C2C}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right)\right), \tag{EC.26}$$

$$\mathbb{E}\left[\left(y_1 - d^{\mathrm{C2C}}\right)^+\right] = \lambda\big(1-q(a)\big)\sum_{i=1}^{T^{\mathrm{R}}} u_i^{C2C}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right). \tag{EC.27}$$

Using the results presented in (EC.20) – (EC.27), we can rewrite (EC.19) as

$$\begin{aligned}\mathbb{E}\left[\hat{\Pi}(a)\right] =& R^{\mathrm{W2C}}\lambda - C^{\mathrm{C2W}}\lambda\sum_{i=1}^{T^{\mathrm{R}}} u_i^{C2W}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right)\\
&- C^{\mathrm{C2W}}\lambda\big(1-q(a)\big)\sum_{i=1}^{T^{\mathrm{R}}} u_i^{C2C}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right)\\
&+ R^{\mathrm{C2C}}(a)\lambda q(a)\sum_{i=1}^{T^{\mathrm{R}}} u_i^{C2C}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right)\\
&- \left(C^{\mathrm{HI}}p^{\mathrm{HI}} + C^{\mathrm{LHI}}(1-p^{\mathrm{HI}}) + C^{\mathrm{C2C}}(a)p^{\mathrm{FR}}\right)\lambda q(a)\sum_{i=1}^{T^{\mathrm{R}}} u_i^{C2C}\prod_{j=1}^{i-1}\left(1-u_j^{C2C}-u_j^{C2W}\right).\end{aligned}$$

Recall that when T^{M}=1, the probability that a C2C returning item is sold within the assignment period is $p^{\mathrm{M}}(a)$=$q(a)$, implying that

$$\begin{aligned}\mathbb{E}\left[\hat{\Pi}(a)\right] =& \lambda\Bigg(R^{\mathrm{W2C}} - C^{\mathrm{C2W}}p^{\mathrm{C2W}} - C^{\mathrm{C2W}}p^{\mathrm{C2C}}\left(1-p^{\mathrm{M}}(a)\right) + R^{\mathrm{C2C}}(a)p^{\mathrm{M}}(a)p^{\mathrm{C2C}}\\
&- \left(\left(C^{\mathrm{HI}}p^{\mathrm{HI}} + C^{\mathrm{LHI}}(1-p^{\mathrm{HI}})\right) + C^{\mathrm{C2C}}(a)p^{\mathrm{FR}}\right)p^{\mathrm{M}}(a)p^{\mathrm{C2C}}\Bigg)\\
=& \lambda\mathbb{E}\left[\Pi_{\mathrm{B}}(a)\right].\end{aligned}$$

□