правительство российской федерации
федеральное государственное автономное образовательное
учреждение высшего профессионального образования
"Национальный исследовательский университет
"Высшая школа экономики"

негосударственное образовательное учреждение
высшего образования
"российская экономическая школа"(институт)

выпускная квалификационная работа

Почему покупатели со временем начинают возвращать больше в онлайн-магазины?

Программа Бакалавр экономики

Совместная программа по экономике НИУ ВШЭ и РЭШ

Автор:
А.Д. Щеткина

Научный руководитель:
Д.И. Дзябура

Москва, 2021 г.

Аннотация

С появлением онлайн-магазинов покупатели стали учитывать возможность возврата при принятии решений о покупках. В этой работе я моделирую покупки и возвраты как две взаимосвязанные части одного процесса принятия решений, вдохновленного моделями поиска, и объясняю динамику этих переменных с помощью двух механизмов. Механизм того, как покупатели узнают о подробностях процесса возврата, объясняет их динамику на уровне популяции: в момент совершения первого возврата, покупатель узнает свой индивидуальный уровень усилий, необходимый, чтобы вернуть товар. После этого или меняется склонность покупателя возвращать товары, или он вообще перестает покупать что-либо у данного продавца. Я показываю, что для широкого семейства распределений такой механизм приводит к растущему агрегированному проценту возвратов. Механизм того, как покупатели знакомятся с брендом, демонстрирует динамику возвратов на индивидуальном уровне. В этой модели у товаров есть общая неизвестная характеристика — качество бренда, и покупатель выбирает оптимальную последовательность покупок, зная о возможности вернуть товар, а также зная, что раскроет часть информации о бренде в процессе. Я численно решаю эту модель и нахожу, что для брендов более высокого класса индивидуальный процент возвратов со временем растет. Модели предлагают различающиеся эмпирические предсказания для брендов более высокого и более низкого классов, и для валидации моделей я тестирую эти предсказания на данных, полученных от крупного российского онлайн-магазина. Групповая и индивидуальная динамика возвратов в данных совпадает с предсказанной теоретически.

Abstract

The emergence of online shopping led customers to take into account the presence of returns when making their purchase decisions. In this paper, I present a theoretical framework that models purchases and returns as two integrated parts of the same search-like process and sheds light on their dynamics. The model of customer learning about the return process explains population-level dynamics of the return rate: the unknown value of each customer's hassle cost of the return is drawn from a common distribution at the moment of the first return incidence. After that, the customer may adjust her return behavior, or stop purchasing from the retailer entirely. For a large family of distributions, this results in an aggregate-level increase in the return rate. The model of learning about the brand explores the individual-level dynamics. In this model, the items sold share a common uncertain attribute — brand quality, and the customer chooses the optimal sequence of purchasing, taking into account the presence of returns and the fact that some of the information of the attributes will be revealed in the process. I solve this model numerically and find that for higher-end brands with high and predictable quality, and low and unpredictable fit the individual return rate increases over time. I test differentiating between higher- and lower-end brands predictions of this framework on the data available from a large Russian online retailer. The group-level and individual-level dynamics are consistent with the models' predictions. Additional evidence on purchase and return rate dynamics also supports the hypothesized mechanisms.

Содержание

1 Introduction **4**

2 Literature Review **5**

3 Theory **9**

3.1 Group Level: Learning about Returns 9

3.2 Individual Level: Learning about the brand 12

4 Simulations **17**

5 Empirical Predictions **19**

6 Estimation **20**

6.1 Data 20

6.2 Main predictions 22

6.3 Additional predictions 26

7 Conclusion **28**

A Proof of Lemma 1 **30**

B Proof of Lemma 2 **31**

C Additional simulations graphs **32**

D Additional summary and descriptive tables **33**

1 Introduction

With the rise of online shopping during the past decade, returns have suddenly become a prominent phenomenon. Back in 2003, Peck and Childers noted the importance of haptic information for customers and predicted the profound impact that the inability of touching an item before purchasing it will have on Internet sales. They hypothesized that for categories in which the "touch-and-feel" aspect is especially important, such as apparel, brick-and-mortar clusters will be integrated to facilitate online retailing. However, instead of pre-purchasing showrooms, the role of revealing the haptic attributes went to post-purchase returns. Now, the average return rate for apparel items sold online (the most returned category) is more than 30 %, which is several times more than in offline stores (*Wall Street Journal*, 2021a). The customers are becoming more and more used to online shopping and to having an opportunity to return. Moreover, online retailing promises to become even more prevalent in the future, as during the pandemic many people developed a habit to purchase a number of different categories and items online, and are planning to continue doing so (*Wall Street Journal*, 2021a). As customers are aware of returns, they incorporate this expectation in their purchasing decisions, so that purchases and returns dynamically influence one another. Still, the existing marketing literature is focused on choice models of purchases, while customer-side models incorporating purchases and returns into a one two-part decision making process are scarce.

To understand how the presence of returns affects purchases and returns themselves is important for firms for several reasons. First, clearly the presence of returns affects companies' profits on its own: 30 % more returned items is 30 % less revenue with the same number of purchases. Moreover, return logistics are costly, and the salvage value of a returned item may be low. It is particularly true for online retailers: since items are not returned in the exact same store that sells them, as is the case with brick-and-mortar retailers, the logistic cost of processing a return might be especially high. In fact, it is sometimes so high that for a company it is optimal to let customers to keep the items they want to return while still getting a refund (*Wall Street Journal*, 2021b). Therefore, it is essential for firms to understand what determines the return rate to be able to accurately predict its profit. It would not be such an important problem though if a firm was able to predict the future return rate of a customer based on a few initial purchases. However, since customers are learning in the process of purchasing and returning, either about return process or about items sold, the initial values of a customer's return rate might not be a good reflection of its future values (El Kihal et al., 2021). Thus, it is particularly important to build a model exploring the dynamics of the return rate. Finally, customers' purchasing patterns are also likely to be affected by the presence of returns, since customers incorporate the knowledge that a return is possible into their purchasing decisions. Understanding these potential effects may help firms adjust their calculations of customer lifetime value, retention strategies, assortment display, and recommendation systems.

I develop a rational choice theoretical framework that combines purchase and return decisions as two parts of the same search-like process and use it to explore the dynamics of purchases and returns. By differentiating between customers' learning about the brand and learning about the return process, I derive the population-level and the individual-level dynamics of return rates that each of these two types of learning produces. Both models imply specific empirical predictions for dynamics of return rates as well as of purchases, the most prominent of them being heterogeneous effects on higher-end versus lower-end brands. I find that learning about the return process results in a population-level increase in the return rates for lower-end brands, but there are no individual effects on this type of brands. In contrast, for higher-end brands, both population and individual return rates are increasing over time. I test these and additional predictions on the data available from a large Russian online retailer, checking the validity of the model.

The rest of the paper is organized as follows: Section 2 provides the review of relevant literature. Section 3 describes the two theoretical models of learning about the return process (group-level) and learning about the brand (individual-level). Section 4 presents the results of simulations studies of the model of learning about the brand. The empirical predictions implied by the two models are described in Section 5. Section 6 presents the empirical part of the work, including data description, estimation, and results. Section 7 concludes.

2 Literature Review

This paper is related to two major streams of literature: to the literature on returns and to theoretical consumer search literature. With the increasing share of online shopping in recent years, the problem of returns is becoming more prominent in business. This problem is reflected in growing marketing literature devoted to the analysis and predictions of returns. In terms of empirical findings, the closest to my work is the working paper of El Kihal, Erdem and Schulze (2021), who document a counterintuitive result that the return rate increases over the course of the relationship between a customer and an online retailer. They attribute this result to the relationship effect, which is reflected in changes of a customer's patterns of purchasing over time, and to the return habit formation, which links greater return rates in past periods to greater future return rates. I complement their findings by developing a theoretical rational choice framework: namely, I propose the mechanism for the implied increase in individual return rates. Moreover, I suggest an alternative group-level explanation for the same phenomenon and argue that both mechanisms contribute to the observed main effect in the way that in some known cases the return rate actually does not change or even decreases over time.

Previous literature suggests that the ultimate cause of relatively higher returns in the context of online retailing is the fact that some item attributes can be evaluated only physically. Peck and Childers (2003) discuss the concept of touch-and-feel attributes and establish their role in product returns. In a recent paper, Dzyabura et al. (2019) confirm that not only is there an inherent uncertainty in some of item characteristics, but also large discrepancies between online and offline product evaluations that are not fully explained by uncertainty. They explain this part of discrepancies by different weights that a customer places on product attributes when evaluating them online and offline. I differ from that work in that I assume online evaluation and returns to be driven by optimal search behavior of customers, as opposed to bias.

Still, to model returns as a part of rational choice model is common practice in literature. To implement the idea of uncertain product attributes in an empirical study of returns, Anderson et al. (2009) develop a general econometric model to study returns under the assumption of unobservable before purchase fit, which follows normal zero-mean distribution. They use this model to estimate a customer's option value of returns and to optimize return policy. This model is extended by Hong and Pavlou (2014), who assume uncertain quality in addition to fit. In their model, fit uncertainty follows Salop's circle model, while quality may have any symmetric distribution. They show that uncertainty both in fit and in quality is linked positively with the return rate. They test it empirically and also demonstrate that product images and reviews decrease the level of uncertainty (this finding is further explored by Dzyabura et al. (2020) who demonstrate that product images improve the accuracy of returns predictions). Similarly to Hong and Pavlou (2014), I assume uncertainty both in fit and in quality to be the cause of a return for an individual level model, however, I impose less restrictions on their distributions, and, more importantly, I study the changes in one customer's return rate and link it to purchase dynamics rather than explain the difference in return behavior between customers or platforms.

A large part of literature concerning returns is devoted to the effects of different return policies: a metaanalysis can be found in Janakiraman et al. (2016) and Minnema et al. (2017). The closest to my work are papers discussing the role of the so-called "effort" cost policies: I use a similar concept of the hassle cost of a return as a pivotal attribute in the group-level model. One example of these studies is Griffis et al. (2012). In this paper, refund speed is used as a proxy for effort return cost. It is shown that those consumers who experienced faster refund speed, start purchasing more frequently, more items per order and items with higher prices. Another example is Bonifield et al. (2010), who demonstrate that lenient policies (including lenient in terms of effort cost policies) can also be a signal of good quality of the retailer. This study describes the effect of minimal return restrictions on a customer's purchase intentions. I differ from these papers in that they are model-free and use data generated in experiments, while I develop a theoretical model and use observational data.

The existing theoretical literature on returns is also primarily concerned with a firm's side rather than with a customer's side. Shulman et al. (2009) derive the optimal restocking fee for a firm in the presence of valuation uncertainty following a Salop's circle type model with uniform utility distribution. Hsiao and Chen (2012) find the optimal restocking fee, assuming Bernoulli distribution of valuation and two types of consumers, whose hassle costs are correlated with valuations. Altug and Aydinliyim (2016) solve for the optimal restocking fee when customers behave strategically, deciding whether to buy an item for a full price or wait for the clearance period. The main distinction of my work is that I focus on a customer's side of the return process and study its dynamics.

There is growing literature similarly exploring a customer's side, namely, potential causes of differences in return rates between different groups of consumers. Petersen and Kumar (2009) show that such factors as whether an item is a gift, multichannel purchasing, purchases within sales periods, etc. may cause changes in the level of the return rate. Narang and Shankar (2019) study the effect of a mobile app introduction on returns. According to their findings, mobile app adopters purchase and return significantly more frequently than those who do not use the app. On an individual level, Shehu et al. (2020) report that free shipping promotions, as opposed to coupons, increase return rates by inducing riskier purchases. They attribute the result to two effects: the first is that free shipping serves as a risk premium for risky purchases, and the second is generation of customers' affect which translates to risk-seeking behavior. I contribute to these studies by exploring the causes of changes in individual return rate in the absence of any changes in a firm's policies.

Both theoretical models that I build for explaining an increase in the return rate are inspired by consumer search literature. Since the process of an online purchase is related to revealing the uncertain value of an item and paying the cost of the return if the attempt is not successful, the analogy to classic search models seems natural. I follow the framework established by Weitzman (1979). In this fundamental paper, an agent solves for the optimal sequence of searching and for a stopping rule. The optimal algorithm is derived via the introduction of the reservation price concept. This general result was extended by several authors in subsequent years. Burdett and Malueg (1981) solve the problem when a customer is searching for several products and cannot return to previously searched alternatives. Carlson and McAfee (1984) explore a similar setting, but with no recall of products searched before. The model is further extended by Gatti (1999). In the latter version of setting, a customer optimizes an indirect utility function within the multi-commodity setting. Necessary and sufficient conditions for the optimality of reservation price type strategies are derived. Ke et al. (2016) solve for optimal strategy with the search of two products when the learning is continuous and the information is never complete. Dzyabura and Hauser (2019) study the role of changing utility in the search process. In their model, a customer updates the weights she places on product attributes when searching.

A number of authors develop various versions of the sequential search model in order to find the market equilibrium. Arbatskaya (2007) demonstrates that in the setting with a predetermined search order, there exists an equilibrium in which a customer with lower search costs finds a better price. In contrast, Branco et al. (2012) show that under the assumption of gradual learning of product attributes, lower search costs can lead to higher prices in equilibrium. Zhou (2014) describes a counterintuitive "joint search effect": when search costs do not depend on the number of items searched within one firm, items sold by the same firm, even with independent valuations, can act as demand complements. My work is particularly close to Liu and Dukes (2013) in that they assume that a product's uncertainty has both a firm's component and an item's component. In this setting, they show that a lesser degree of differentiation between firms may lead to larger product assortment. The distinctive features of my model in comparison to the previous literature include the following: a customer searches for several items but needs not to buy any predetermined number of products, and consequently, needs not to have a stopping rule; information is revealed instantaneously and different alternatives have correlated information; search cost is paid not always but only in case of a low value realization. All these features are necessary to adapt a general search model to a more specific context of online returns.

The classic theoretical search models described above have recently received attention in applications. Ursu (2018) uses the search model to quantify the effect of a position in rankings. The model shows that rankings affect search costs but not purchases if controlled for search. Ursu and Dzyabura (2019) apply the model to the product location problem and find that a retailer has incentives to prioritize a product with lower utility, since it implies the willingness of a customer to search it only at a lower cost. Chen and Yao (2017) explore the value of refinement tools in the search process. They show that refinement tools enhance the amount of search but may hurt consumers in case they are unaware of default rankings by popularity or quality. In a work in progress, Jerath (2021) studies the relationship between consumer pre-purchase search and returns from the point of view of the firm. This model demonstrates that for the firm it might be optimal to set zero restocking fees but introduce additional complications to the return process, such as original package requirement. In contrast to this work, I do not model pre-purchase search (in narrower sense) and returns as two separate but connected phenomena, but I consider returns to be the result of consumer search (in broader sense) itself.

3 Theory

In this section I describe two rational choice models that may potentially explain the increase in the return rate over time. The first model operates on a group level and is related to customers' learning about the return process. The second model shows that an increase in each individual return rate is also compatible with optimal search behavior and is linked to customers' learning about the brand.

3.1 Group Level: Learning about Returns

Suppose a firm sells multiple items to multiple customers. The valuations of the items are uncertain for the customers, however, for all customers and all items valuations are independent and identically distributed with probability density function g_v, that is, a purchase of one item does not reduce uncertainty regarding other items' valuations. Each period each customer decides whether to buy a new item with valuation $v \sim g_v$. g_v is normalized such that the price of the item is taken into account.

In addition to valuation uncertainty, each customer faces uncertainty regarding his hassle cost of the return: intuitively, it is unclear how many forms it will take to fill out to process the return, how far away the customer will need to go to send the returned item back, whether the original packaging is required, and how quickly he will receive the refund. To capture this uncertainty, I introduce the distribution of the hassle cost of the return $c \sim g_c$, which is known to customers. Before a customer has made his first return, he knows the distribution of the hassle cost but not its exact value. At the moment of the first return, the customer's exact value of the hassle cost is revealed, that is, drawn from the distribution g_c. Starting from this point, the customer knows his exact value of the hassle cost.

Since the information available for a customer is different in two different states (before and after the first return is made), I describe and solve a customer's problem separately for these two states:

Before the first return is made

These are the time periods starting from the first purchase until the purchase at which the first return is made. A customer knows the distributions g_v and g_c but not the exact value of the hassle cost. The timeline of each purchase is as follows:

1. The customer decides whether to buy an item in this period (the valuation and the hassle cost are not known).

2. If the customer decides to buy the item, he receives it and the valuation of the item is realized.

3. The customer decides whether to return or not (the valuation is known, the hassle cost is not).
4. If the customer decides to return the item, the hassle cost is realized, the customer pays it and gets rid of the item. In that case the state "after the first return" begins. Otherwise, no new information is realized, and the customer goes through this timeline in the next period.

Note that since the customer does not get any new information between periods, the problem is identical in each period in this state. The customer solves the problem by backward induction. The last decision is whether to return an item with the realized valuation v. In case if the customer keeps it, he will get the utility v. In case the customer returns it, he will have to pay the hassle cost drawn from the distribution g_c. Since v is known by this time, the customer behaves optimally if he returns the item if and only if

$$v < -\mathbb{E}c = -\bar{c} \tag{1}$$

The expected utility from such a purchase and the probability of return (return rate) are hence

$$\mathbb{E}U_{before} = -\bar{c}\,\mathrm{P}(v < -\bar{c}) + \int_{-\bar{c}}^{+\infty} x g_v(x)dx \tag{2}$$

$$RR_{before} = \mathrm{P}(v < -\bar{c}) \tag{3}$$

Taking this into account, the customer decides whether he wants to purchase an item or not (step 1 of the timeline). Optimally, he purchases the item if and only if the expected utility from it is higher than zero:

$$\mathbb{E}U_{before} = -\bar{c}\,\mathrm{P}(v < -\bar{c}) + \int_{-\bar{c}}^{+\infty} x g_v(x)dx \geq 0$$

Note that the expected utility from a purchase is exactly the same in all periods in the "before" state. That is, if it is optimal to purchase in the very first period (optimal to purchase at all), it is also optimal to purchase in all the next periods until the hassle cost is realized.

Also note that the customer behaves in the "before" state as if his prior belief on his value of the hassle cost is $\bar{c}$. In fact, the customer knows and takes into account the entire distribution of the hassle cost, but only its mean is relevant for his decisions.

After the first return is made

Once the valuation of the item dropped lower than $-\bar{c}$, as explained above, the customer returns the item and learns his true hassle cost, c_{after}. For all the periods starting from now on the timeline of the decisions is as follows:

1. The customer decides whether to buy an item in this period (the valuation is not known, the hassle cost is).
2. If the customer decides to buy the item, he receives it and the valuation of the item is realized.
3. The customer decides whether to return or not (the valuation and the hassle cost are known).
4. If the customer decides to return the item, the customer pays the known hassle cost and gets rid of the item.

Again solving by backward induction, on step 3 the customer returns the item if and only if the realized value is lower than paying the known hassle cost:

$$v < -c_{after} \tag{4}$$

The expected utility from such a purchase and the return rate are:

$$\mathbb{E}U_{after} = -c_{after}\,\mathrm{P}(v < -c_{after}) + \int_{-c_{after}}^{+\infty} x g_v(x)dx \tag{5}$$

$$RR_{after} = \mathrm{P}(v < -c_{after}) \tag{6}$$

On step 1, the customer purchases an item if and only if

$$\mathbb{E}U_{after} = -c_{after}\,\mathrm{P}(v < -c_{after}) + \int_{-c_{after}}^{+\infty} x g_v(x)dx \geq 0 \tag{7}$$

Depending on the realized value of the hassle cost, there are several possible scenarios. If the realized value is lower than the mean of the distribution, the customer starts returning more frequently:

$$c_{after} < \bar{c} \Rightarrow RR_{after} = \mathrm{P}(v < -c_{after}) > RR_{before} \tag{8}$$

If the realized value is higher than the mean of the distribution, the customer starts returning less frequently:

$$c_{after} > \bar{c} \Rightarrow RR_{after} = \mathrm{P}(v < -c_{after}) < RR_{before} \tag{9}$$

However, there are also customers for whom the realized value of the hassle cost is so high that for them it is optimal to leave the market.

Lemma 1 *For any given distribution of valuation g_v if $\mathbb{E}v < 0$ and $\mathbb{E}U_{before} \geq 0$ there exists a unique critical value of the realized hassle cost of the return $\tilde{c} \geq \bar{c}$, such that the customer is indifferent between purchasing items in the future or not. That is, the customer will leave the market after the first return if and only if the realized hassle cost $c_{after} \geq \tilde{c}$.*

Proof. See Appendix A. ■

The assumption that $\mathbb{E}v < 0$, while might not be correct, is plausible. Recall that valuation already includes price paid, so this assumption can intuitively be formulated as "if any returns are prohibited, a customer would not want to purchase from the online retailer". This assumption makes sense especially for apparel items because of the combination of three factors: prices are relatively high, the level of uncertainty (e.g., size uncertainty) when purchasing online is relatively high, finally, if the size does not fit, the item is of very little use, so its net valuation is close to minus paid price. Recall that the second assumption $\mathbb{E}U_{before} \geq 0$ is purely technical — it simply means that it is optimal for a customer to purchase anything at all from the retailer. Otherwise, we do not observe such a customer in the purchase data.

That is, after the true hassle cost is revealed, the customers are divided into three groups:

1. Those who start returning more frequently. The size of this group is $\mathrm{P}(c_{after} < \bar{c})$
2. Those who start returning less frequently. The size of this group is $\mathrm{P}(\bar{c} < c_{after} < \tilde{c})$
3. Those who leave the market and are not observed after the first return. The size of this group is $\mathrm{P}(\tilde{c} < c_{after})$

Since the third group is eventually left out of observation, the dynamics of the return rate is determined by the relative sizes of the first two groups. In close to symmetric distributions of the hassle cost g_c and in cases when the critical value of the hassle cost is close to the mean, the first group is larger, so the increasing effect dominates, and we observe the overall increasing pattern of the return rate (see Figure 1). This effect is on a group level, since it comes from heterogeneity in the hassle cost among customers and since the aggregate-level return rate increases, while it is not the case for all individual return rates.

3.2 Individual Level: Learning about the brand

Consider a firm selling two items under the same brand name to a customer in two periods. One of the items is "safe" — with a lower level of uncertainty (such as a plain white t-shirt), and the other is risky — with a higher level of uncertainty (such as an elaborate red dress). Following Hong and Pavlou (2014), the items have two attributes: fit (unique

Three groups after the realization of the hassle cost

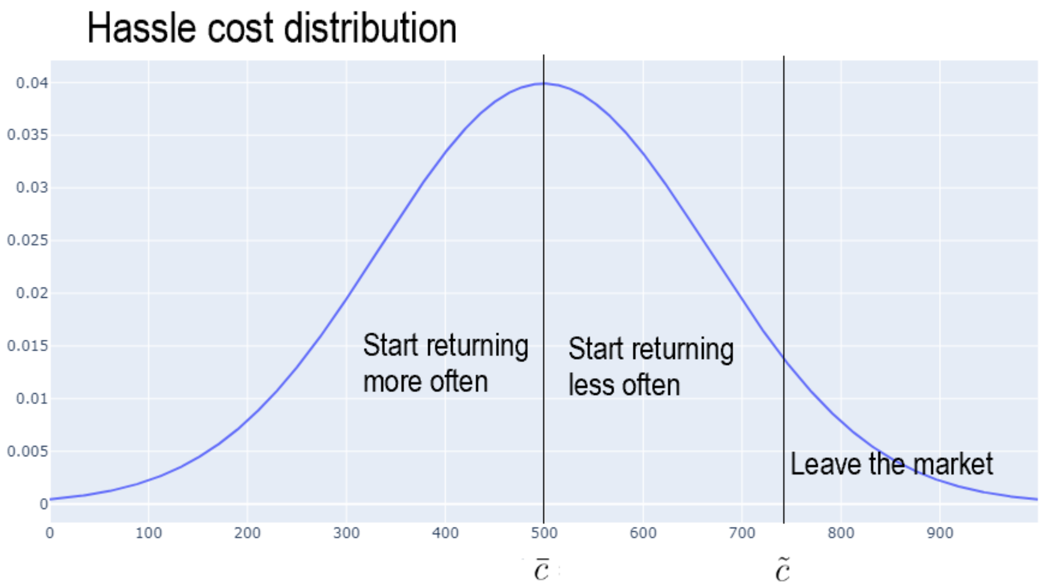

Figure 1: This Figure shows the relative size of the three groups of customers after the realization of the hassle cost. $\bar{c}$ denotes the mean of the distribution, and $\tilde{c}$ denotes the critical value of the realized hassle cost beyond which a customer leaves the market.

for each item) and quality (common for the brand). To capture a safe/risky relationship, let the customer's valuation of the safe item to depend only on quality, and the customer's valuation of the risky item to depend both on quality and on its fit. Specifically, in the simplest case:

$$\begin{cases} v_{safe} = q \\ v_{risky} = f + q, \end{cases} \tag{10}$$

where v_{safe} is the customer's valuation of the safe item, v_{risky} is the customer's valuation of the risky item, q is the quality of the brand and is the same for the two items, and f is the fit of the risky item. In other words, the fit of the safe item is known for sure, and the quality is such that this known fit of the safe item is normalized to zero.

Both fit and quality are inherently uncertain at the time of purchasing an item from the website: think of touch-and-feel attributes of fabric that constitute a part of uncertainty about the quality, and of the size which is an important and uncertain part of the fit. Because of this uncertainty, I allow both fit and quality to follow their respective and independent general distributions, with respective probability (cumulative) density functions g_f (G_f) and g_q (G_q). Let them be normalized such that the prices of the items are already included in the customer's valuations, v_{safe} and v_{risky}. Once the item is purchased, the true values of all its attributes are revealed. Also let $c \geq 0$ be the cost of the return of one item (in form of a restocking fee and/or the hassle cost) for the customer.

This value is constant and known to the customer, and there is no economy of scale from returning two items at the same period.

Table 1: Description of all variables in the individual-level model

q	quality (same for both items)
g_q	probability density fuction of quality
G_q	cumulative density function of quality
f	fit (of the risky item)
g_f	probability density function of fit
G_f	cumulative density function of fit
c	return cost
v_{safe}	valuation of the safe item, $v_{safe} = q$
v_{risky}	valuation of the risky item, $v_{risky} = q + f$

The customer solves a dynamic search problem. If she purchases the safe item in the first period, by the beginning of the second period she will know the true quality of the risky item, but not its fit. Vice versa, if the risky item is purchased in the first period, by the beginning of the second period the customer will know the true quality and hence her valuation of the safe item. Taking these future updates of the information available to her into account, the customer solves for the optimal sequence of purchasing, in other words, whether to buy the safe item in the first period and then to decide whether to buy the risky in the second period, or vice versa. Note that it is never strictly optimal to purchase both items in the first period, since by delaying the purchase of one of the items to the second period, a customer learns about its quality, and either still purchases the item without any loss or does not purchase the item because the quality is too low — in this case she is strictly better off from waiting. Therefore, the customer compares the sums of her expected utilities in two periods in two cases (assuming it is optimal to buy anything at all): when the safe item is purchased in the first period or when the risky item is purchased in the first period.

The safe item is purchased first

Suppose the safe item is purchased in the first period. Its valuation is revealed, hence the true quality of the risky item is revealed. The customer returns the safe item if its realized valuation, i.e. quality, is lower than the opportunity cost of paying the return cost ($-c$). That is, with probability $\mathrm{P}(q < -c)$, the customer gets $-c$. If the realized valuation is higher than paying the return cost, the customer gets q according to the probability density function g_q. The expected utility from purchasing the safe item in the first period is therefore

$$\mathbb{E}U_1^{safe} = -c\,\mathrm{P}(q < -c) + \int_{-c}^{+\infty} x g_q(x) dx \tag{11}$$

In the second period, the customer decides whether to purchase the risky item, knowing its true quality.

Lemma 2 *For any given distribution of fit g_f and return cost c there exists a unique critical value of quality $\tilde{q}$, such that the customer is indifferent between purchasing the risky item or not in the second period. That is, the risky item will be purchased if and only if the realized in the first period $q \geq \tilde{q}$.*

Proof. See Appendix B. ■

That is, even before the first period the customer knows her critical value $\tilde{q}$ and that if the realized value of q in the first period is higher than this $\tilde{q}$, she will purchase the risky item in the second period. Otherwise, she will not purchase the risky item in the second period. Again, the item is returned if and only if its realized valuation ($v_{risky} = f + q$) is higher than paying the return cost. Thus, from the point of view of the first period, when the quality is not revealed yet, the expected utility from purchasing the risky item in the second period is the following:

$$\mathbb{E}U_2^{risky} = \int_{\tilde{q}}^{+\infty} g_q(x) \left(\int_{-c-x}^{+\infty} (x+y) g_f(y) dy \right) dx - c \int_{\tilde{q}}^{+\infty} g_q(x) G_f(-c-x) dx \qquad (12)$$

The first term represents the utility from keeping the item in case if $q > \tilde{q}$ (otherwise, the item is not purchased) and $f + q > -c$ — or $f > -c - q$ (otherwise the item is returned). The second term stands for paying the return cost in case if both $q > \tilde{q}$, so the customer purchases an item, and the realized value of fit is lower than $-c - q$, so the customer returns the item.

The total expected utility from this sequence of purchases is:

$$\begin{aligned}\mathbb{E}U_{safe\ first} = & \\ & - c\,\mathrm{P}(q < -c) + \int_{-c}^{+\infty} x g_q(x) dx + \\ & \int_{\tilde{q}}^{+\infty} g_q(x) \left(\int_{-c-x}^{+\infty} (x+y) g_f(y) dy \right) dx - c \int_{\tilde{q}}^{+\infty} g_q(x) G_f(-c-x) dx\end{aligned}$$

The return rate in the first period in this sequence is

$$RR_1^{safe} = \mathrm{P}(q < -c) \qquad (13)$$

The return rate in the second period in this sequence is

$$RR_2^{risky} = \int_{\tilde{q}}^{+\infty} g_q(x) G_f(-c-x) dx \qquad (14)$$

The risky item is purchased first

Suppose the risky item is purchased in the first period. Its valuation is revealed, hence the true quality and the true fit are revealed, so that the valuation of the safe item becomes fully known. The customer returns the risky item if its realized valuation (the sum of quality and fit) is lower than the opportunity cost of paying the return cost ($-c$). The expected utility from the risky purchase in the first period is therefore similar to the equation 12, except for quality that is no more lower bounded by its critical value:

$$\mathbb{E}U_1^{risky} = \int_{-\infty}^{+\infty} g_q(x) \left(\int_{-c-x}^{+\infty} (x+y) g_f(y) dy \right) dx - c \int_{-\infty}^{+\infty} g_q(x) G_f(-c-x) dx \quad (15)$$

In the beginning of the second period, the customer knows exactly her valuation of the safe item, which is equal to the realized quality, so she purchases it if and only if the realized quality is greater than zero:

$$\mathbb{E}U_2^{safe} = \int_0^{+\infty} x g_q(x) dx \quad (16)$$

The total expected utility from this sequence of purchases is:

$$\mathbb{E}U_{risky\ first} = \\ \int_{-\infty}^{+\infty} g_q(x) \left(\int_{-c-x}^{+\infty} (x+y) g_f(y) dy \right) dx - c \int_{-\infty}^{+\infty} g_q(x) G_f(-c-x) dx + \\ \int_0^{+\infty} x g_q(x) dx$$

The return rate in the first period in this sequence is

$$RR_1^{risky} = \int_{-\infty}^{+\infty} g_q(x) G_f(-c-x) dx \quad (17)$$

The return rate in the second period in this sequence is 0, since the valuation of the safe item is fully known.

$$RR_2^{safe} = 0 \quad (18)$$

The customer chooses the sequence with the highest total expected utility (if this is more than zero). Note that a necessary condition for the return rate to increase from the first to the second period is that a customer purchases the safe item first. Otherwise, as shown above, the return rate in the second period is zero. The customer purchases the safe item in the first period if and only if $\mathbb{E}U_{safe\ first} > \mathbb{E}U_{risky\ first}$, or, after canceling out, if

$$
\begin{aligned}
&- c\,\mathrm{P}(q < -c) + \int_{-c}^{0} x g_q(x) dx > \\
&\int_{-\infty}^{\tilde{q}} g_q(x) \left(\int_{-c-x}^{+\infty} (x+y) g_f(y) dy \right) dx - c \int_{-\infty}^{\tilde{q}} g_q(x) G_f(-c-x) dx
\end{aligned}
$$

However, while this condition is necessary, it is not sufficient. Depending on parameter values and distributions, this inequality may or may not hold, as well as the inequality $RR_1^{safe} < RR_2^{safe}$. I numerically solve the customer's problem for different parameter values, describe the characteristics of relevant combinations of them, and provide intuition in section 4.

4 Simulations

Since the solution of the model with the safe and the risky items depends on the distributions and parameter values, I solve the model numerically for different combinations of parameter values. I set the fit and the quality to be normally distributed, which results in 5 parameters to choose: quality mean, quality variance, fit mean, fit variance, and the cost of the return. The grid of the parameters is as follows:

$$
\mu_f, \mu_q \in \{-10, -8, -6, -4, -2, 0, 2, 4, 6, 8, 10\} \tag{19}
$$

$$
\sigma_f, \sigma_q \in \{1, 3, 5, 7, 9, 11, 13, 15, 17, 19\} \tag{20}
$$

$$
c \in \{1, 3, 5, 7, 9, 11, 13, 15, 17, 19\} \tag{21}
$$

In total it accounted for 121,000 possible parameter combinations. I numerically solve the model for each of them. 7,146 out of 121,000 combinations produced the increase in the return rate from the first period to the second (recall that a necessary condition is that the safe item is purchased in the first period). Since the combinations of parameters are in five-dimensional space, I plot them pairwise, with marker size denoting the number of combinations of other three parameters that together with the plotted two produce the increasing pattern in the return rate. Figure 2 presents the values of fit and quality means that result in the increasing return rate, and Figure 3 presents the fit and quality variances. Additional graphs of the simulations results can be found in Appendix C.

The results suggest that the individual return rate should increase for brands that have high and predictable quality and low and unpredictable fit. This corresponds to a common perception of higher-end brands: a customer is almost certain that the fabric and overall quality of any item from such a brand is good enough, while the complexity of design of dresses or suits from such brands prevents them to be a good fit for everyone.

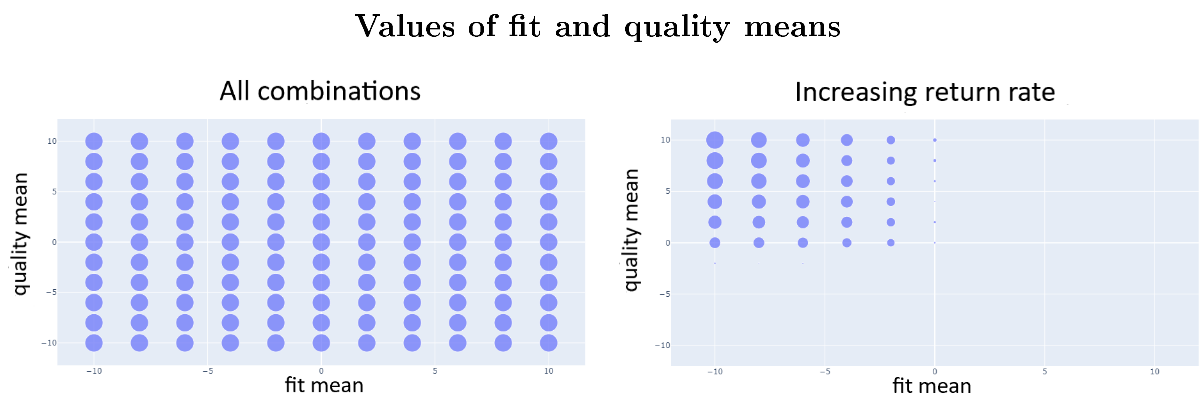

Figure 2: On the left-hand side: all combinations of fit and quality means that are solved. On the right hand side: combinations of fit and quality means that result in the increasing pattern in return rate. The marker size corresponds to the number of combinations of these two parameters with the other three that lead to the increasing dynamics.

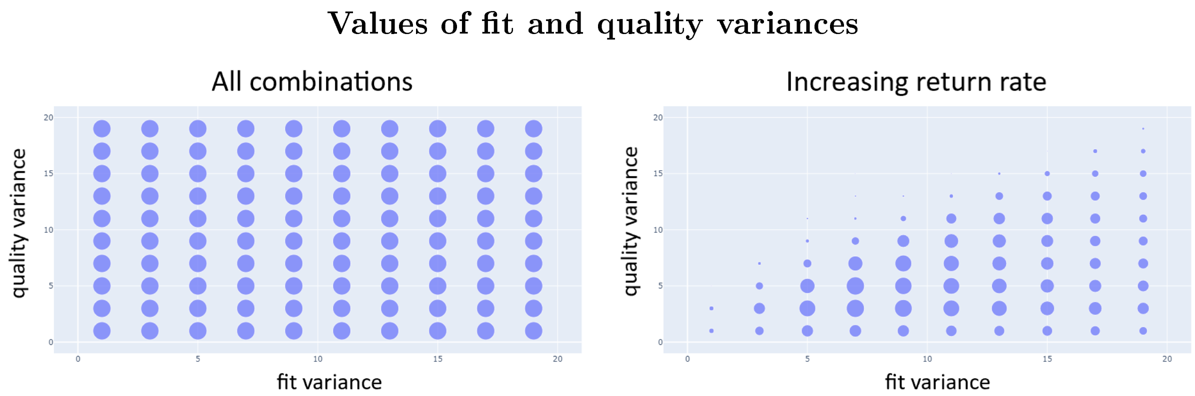

Figure 3: On the left-hand side: all combinations of fit and quality variances that are solved. On the right hand side: combinations of fit and quality variances that result in the increasing pattern in return rate. The marker size corresponds to the number of combinations of these two parameters with the other three that lead to the increasing dynamics.

The intuition of this result is the following. Since the quality is high and predictable, it means that the valuation of the safe item is also high and predictable, while the valuation of the risky item is not. Since the valuation of the safe item is high, the customer is indifferent between purchasing it in the first or in the second period: the value of any new information on this item is insignificant. However, any new information about the risky item is very helpful in making the decision, since its valuation is not as high and not as predictable. Thus, it is optimal to purchase the safe item first, learn the exact value of the quality, and only then decide whether to purchase the risky item or not. Since the fit is still highly unpredictable, the return rate increases in the second period.

5 Empirical Predictions

Since the individual-level mechanism differentiates between higher- and lower-end brands, the empirical predictions for these two types of brands would be different. Recall that for lower-end brands only the group-level mechanism should work, whereas for higher-end brands both mechanisms should contribute to the dynamics of return rate. Since both these mechanisms increase the return rate over time, the first prediction is consistent with El Kihal et al. (2021) — note that this relationship is purely descriptive:

H 1 *For both higher- and lower-end brands the average return rate increases with the number of order.*

However, since the effect of the group-level mechanism comes from customers' heterogeneity and not from common increasing trend, this effect should disappear once we control for it. For lower-end brands the group-level mechanism is the only one driving the effect, hence the next prediction:

H 2 *For lower-end brands the average return rate does not increase with the number of order once controlled for customer fixed effects.*

For higher-end brands though the individual-level mechanism also works, which by definition creates a common increasing trend for customers. For such brands, the main effect should persist even after controlling for customer fixed effects:

H 3 *For higher-end brands the average return rate increases with the number of order once controlled for customer fixed effects.*

Formulated above are the main predictions of the two models. In addition to them, more specific predictions can be suggested. The individual-level model implies that within higher-end brands, customers start with safe items and then move to risky. This model also implies that if a customer started with purchasing a risky item, the return rate of such a customer should not increase over time, as the customer learns the valuation of the safe item.

H 4 *For higher-end brands the share of risky items purchased increases with the number of order.*

H 5 *If a customer has purchased a risky item in the first period, her return rate does not increase with the number of order.*

The group-level model also implies its specific predictions. Griffis et al. (2012) show that the customers exposed to a lower hassle cost of the return start purchasing more frequently. Since according to the model the customers with the highest realized hassle cost leave the market over time, we should observe that frequency of purchases increases with the number of order. Griffis et al. (2012) also demonstrate that low hassle cost of the return leads to an increased number of items per order. They attribute this effect to building loyalty to a retailer who provides a low hassle cost of returns. However, for retailers with typically high frequency of orders and diverse set of product categories, the loyalty effect may not be translated directly in more items per order purchased. In fact, the customers with the lowest hassle cost of the return may find it optimal to purchase more frequently and fewer items per order: if the hassle cost is high, it is optimal to wait until a large enough shopping cart is collected and return several items at a time to gain economy from scale. If the hassle cost is low, a customer may place an order anytime she needs something, not waiting her shopping cart to pile up. In that case the building loyalty may be expressed in more items bought per week (as opposed to per order), so that a decrease in basket size is offset by an increase in frequency. Similarly to frequency, the group-level model suggests that in that case the number of items per order decreases with the number of order, but the number of items purchased per week increases with the number of week.

H 6 *Frequency of purchases increases with the number of order.*

H 7 *The number of items purchased per order decreases with the number of order, while the number of items purchased per week increases with the number of week.*

6 Estimation

6.1 Data

I use data from a large Russian multi-category online retailer selling food, FMCG, household and beauty items, multi-brand apparel, etc. The most popular categories of purchases and returns can be found in Tables 8 and 9 in section D of Appendix. The data is available on all purchases made by 4,997 customers from January 1, 2019 to December 31, 2019, which in total provides us with 625,127 customer-order-item instances, 62,005 of which are apparel. 5,256 of these items are returned; 2,760 of them are apparel.

In the estimation, I often restrict samples to apparel purchases for two reasons. Firstly, as mentioned above, is that apparel provides a natural venue to study returns, as valuation of an apparel item is highly uncertain, and its price is relatively high. This is supported by the fact that apparel is only the sixth most popular category in terms of purchases, accounting for less than 10% of total, while it is the most popular category in terms of returns, accounting for more than 50% of them. The second reason is that more specific data, such as brand and title, is available for apparel items, which allows me to identify high versus low quality brands, and riskier and safer categories. More precise information is available for 43,165 out of 62,005 purchased apparel items (for 1,724 out of 2,760 returned).

The distributions of orders per customer and money spent on the platform during a year have mainly intersecting sets of outliers to the right. Since such customers are likely to be coming from a different data generating process, I remove them on the data preprocessing step, which leaves me with 4,919 customers and 570,044 customer-item-order observations. Summary statistics are presented in Table 2.

Table 2: Summary Statistics

	Mean	SD	p25	p50	p75
		A. Customer-level			
Days between orders	12.1	21.2	1.7	5.2	13.5
Number of items per order	5.2	6.1	1	3	7
Rubles per order	2169.3	3067.7	590.0	1335.8	2722.7
Number of orders made	22.1	21.0	8	15	28
Registration date	2015/09/10	–	2013/06/20	2017/02/15	2018/11/26
		B. Item-level (all categories)			
Full price	675.1	1838.1	117	252	599
Price discount	0.23	0.18	0.08	0.21	0.34
Paid price	414.5	1140.7	79.2	165.8	378.3
Delivery cost	41.9	123.5	0	0	0
Delivered by courrier	0.55	0.50	0	1	1
Delivered to a staffed pickup point	0.37	0.48	0	0	1
Delivered to an automated pickup point	0.08	0.27	0	0	0
Returned	0.008	0.089	0	0	0
		C. Item-level (apparel purchases)			
Full price	1523.9	2868.6	325	697	1690
Price discount	0.36	0.22	0.20	0.38	0.53
Paid price	718.9	1058.3	180	355.5	832.5
Delivery cost	37.5	110	0	0	0
Delivered by courrier	0.40	0.49	0	0	1
Delivered to a staffed pickup point	0.50	0.49	0	1	1
Delivered to an automated pickup point	0.09	0.29	0	0	0
Returned	0.04	0.20	0	0	0

6.2 Main predictions

To test hypotheses H1 – H3, I estimate the linear probability model of return incidence on the number of order (following El Kihal et al., 2021) for four different samples: total sample of purchases (Table 3), apparel purchases (Table 4), lower-end brands apparel purchases (Table 5), and higher-end brands apparel purchases (Table 6). Subsamples of higher- and lower-end brands can be found in Table 10 in the Appendix. The linear probability model is used to include customer and calendar week fixed effects.

The results of specifications (1) – (3) in all tables (univariate OLS, OLS with covariates, and OLS with covariates and calendar week fixed effects — all without customer fixed effects) provide evidence in support of H1: the return rate increases with the number of order. Moreover, as is consistent with the models, the largest effect is observed for the subsample of high quality brands: each ten additional orders increase the return rate by 1.32 percentage points. It is expected, since for such brands both group-level and individual-level mechanisms contribute to an increase, as opposed to only the former for low quality brands. Still, the estimated effect is arguably small relative to the findings of El Kihal et al. (2021). This difference can be attributed to the fact that in Russia the culture of online purchases and returns is still nascent, as opposed to European countries.

The hypothesis H2 is supported by the results of specifications (4) – (6) (univariate OLS, OLS with covariates, and OLS with covariates and calendar week fixed effects — all including customer fixed effects) in Tables 3 – 5. For low quality brands, the effect of the number of order on return rate becomes insignificant, just as implied by the models. Note that this is also the case for all apparel items and for the total sample: this is probably reflecting the fact that the platform mostly sells lower-end apparel brands, and in general is not considered to be a venue for shopping for premium clothing items.

Finally, columns (4) and (5) of Table 6 are suggestive in favor of H3. Higher-end brands are the only sample of the four that preserve the effect of interest after including customer fixed effects, which is in line with the individual-level model. In the specification (6), however, the effect becomes insignificant. A plausible explanation to it is that including customer and calendar week fixed effects along with order- and item-level covariates leaves too little remaining variation on the sample of 2,629 observations to credibly estimate the effect of interest. Moreover, the effect of interest is obviously time-related, so it is even more probable that it would be absorbed by week fixed effects on small samples. The fact that the inclusion of week fixed effects did not change much the estimates in any other specification provides the reasoning to believe that specifications (4) and (5) are credible. In addition, since higher-end brands are the smallest sample of the four (because they are not common for the platform) and yet the only one to pick up positive and significant effect of the number of order on the return rate, it is more convincing that H3 is supported by the data.

Table 3: Probability of returning by the number of order: total sample

	(1)	(2)	(3)	(4)	(5)	(6)
Number of order	.0048***	.0054***	.0087***	.0007	-.0018	-.0018
	(.0007)	(.0008)	(.0009)	(.0014)	(.0153)	(.0233)
Constant	0.71***	-.31	-.699			
	(.016)	(.66)	(.674)			
Price (per 100 rubles)		.061***	.061***		.055***	.055***
		(.004)	(.004)		(.004)	(.004)
Price discount (1 pp)		2.7***	2.7***		1.94***	2.00***
		(.091)	(.094)		(.142)	(.147)
Quantity ordered		-.018***	-.019***		-.010	-.010
		(.007)	(.007)		(.009)	(.009)
Delivery cost (per 100 rubles)		-.131***	-.138***		-.095***	-.066***
		(.008)	(.010)		(.017)	(.020)
Basket size		-.019***	-.019***		-.017***	-.016***
		(.0009)	(.0009)		(.0024)	(.0025)
Non-premium account		.336***	.400***		.235***	.258***
		(.028)	(.030)		(.071)	(.082)
Prepayment		-.252***	-.255***		-.359***	-.319***
		(.030)	(.030)		(.070)	(.070)
Courrier delivery		.426	.512		.606	.696
		(.661)	(.661)		(.710)	(.710)
Staffed pickup point delivery		.548	.623		.421	.523
		(.661)	(.662)		(.708)	(.708)
Auto pickup point delivery		.302	.368		.606	.684
		(.663)	(.663)		(.709)	(.710)
Customer fixed effects	No	No	No	Yes	Yes	Yes
Calendar week fixed effects	No	No	Yes	No	No	Yes
Number of observations	570,044	570,044	570,044	570,044	570,044	570,044

All specifications are done assuming linear probability model

Dependent variable in percentage points

Clustered on customer level standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4: Probability of returning by the number of order: apparel purchases

	(1)	(2)	(3)	(4)	(5)	(6)
Number of order	.0646***	.0496***	.0743***	-.0112	-.0156	-.0163
	(.0058)	(.0064)	(.0069)	(.0114)	(.0118)	(.0167)
Constant	3.30***	-1.64	-4.19**			
	(.110)	(1.77)	(1.96)			
Price (per 100 rubles)		.326***	.324***		.251***	.251***
		(.015)	(.015)		(.018)	(.018)
Price discount (1 pp)		3.70***	3.96***		1.44***	1.46***
		(.388)	(.403)		(.501)	(.511)
Quantity ordered		-.454***	-.413***		-.151**	-.139*
		(.045)	(.044)		(.077)	(.077)
Delivery cost (per 100 rubles)		-.451***	-.869***		-.125	-.092
		(.060)	(.087)		(.098)	(.162)
Basket size		-.027***	-.029***		-.029*	-.026*
		(.008)	(.009)		(.015)	(.015)
Non-premium account		.394*	1.164***		.777	.977*
		(.214)	(.224)		(.506)	(.545)
Prepayment		.161	.015		-.107	-.051
		(.189)	(.192)		(.465)	(.470)
Courrier delivery		1.55	1.80		2.46	2.64
		(1.75)	(1.76)		(2.22)	(2.22)
Staffed pickup point delivery		2.48	1.76		1.59	1.78
		(1.75)	(1.76)		(2.20)	(2.21)
Auto pickup point delivery		1.39	1.48		1.57	1.69
		(1.76)	(1.77)		(2.23)	(2.24)
Customer fixed effects	No	No	No	Yes	Yes	Yes
Calendar week fixed effects	No	No	Yes	No	No	Yes
Number of observations	57,902	57,902	57,902	57,902	57,902	57,902

All specifications are done assuming linear probability model

Dependent variable in percentage points

Clustered on customer level standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5: Probability of returning by the number of order: lower-end apparel brands

	(1)	(2)	(3)	(4)	(5)	(6)
Number of order	.0808***	.0662***	.0861***	.0262	.0344	.0496
	(.0231)	(.0249)	(.0258)	(.0478)	(.0502)	(.0671)
Constant	5.31***	.492	-1.95			
	(.408)	(2.00)	(5.97)			
Price (per 100 rubles)		1.07***	1.09***		.824***	.841***
		(.134)	(.133)		(.186)	(.187)
Price discount (1 pp)		-3.78***	-2.43		-5.18**	-5.39***
		(1.42)	(1.60)		(1.89)	(1.98)
Quantity ordered		-.121	-.062		.730	.207
		(.106)	(1.08)		(1.31)	(1.39)
Delivery cost (per 100 rubles)		-1.07***	-1.74***		-.526	-.668
		(.254)	(.342)		(.370)	(.537)
Basket size		-.028	-.036		.109*	.109
		(.035)	(.037)		(.060)	(.069)
Non-premium account		.620	1.33		2.10	1.90
		(.909)	(.913)		(2.88)	(3.03)
Prepayment		-.360	-.501		-1.88	-1.46
		(.741)	(.745)		(1.75)	(1.96)
Courrier delivery		-.508	.135		-1.12	-4.83
		(1.07)	(1.09)		(2.95)	(3.05)
Staffed pickup point delivery		2.62***	2.80***		-1.38	-7.91
		(.966)	(.964)		(2.48)	(2.45)
Customer fixed effects	No	No	No	Yes	Yes	Yes
Calendar week fixed effects	No	No	Yes	No	No	Yes
Number of observations	5,780	5,780	5,780	5,780	5,780	5,780

All specifications are done assuming linear probability model

Dependent variable in percentage points

Clustered on customer level standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: Probability of returning by the number of order: higher-end apparel brands

	(1)	(2)	(3)	(4)	(5)	(6)
Number of order	.1123***	.1044***	.1320***	.1414**	.1188*	.0504
	(.0404)	(.0396)	(.0430)	(.0597)	(.0679)	(.1180)
Constant	8.26***	7.54**	1.03			
	(.829)	(3.44)	(3.72)			
Price (per 100 rubles)		.305***	.299***		.174	.206*
		(.051)	(.053)		(.109)	(.114)
Price discount (1 pp)		4.51*	3.79		-3.19	-.547
		(2.50)	(2.72)		(4.90)	(4.99)
Quantity ordered		-5.39***	-5.06***		-.243	1.18
		(.949)	(1.01)		(.985)	(1.05)
Delivery cost (per 100 rubles)		-1.84***	-3.33***		-1.11	.011
		(.543)	(.750)		(.748)	(1.12)
Basket size		.301***	.258***		.175	.185
		(.070)	(.094)		(.131)	(.156)
Non-premium account		1.84	2.02		.416	1.96
		(1.28)	(1.40)		(4.75)	(4.53)
Prepayment		2.24*	1.12		-3.79	-3.98
		(1.25)	(1.27)		(4.62)	(4.32)
Courrier delivery		-7.23***	-7.24***		-4.54	-6.90
		(2.73)	(2.75)		(4.63)	(4.44)
Staffed pickup point delivery		-2.18	-3.09		-10.77*	-10.59**
		(2.77)	(2.75)		(5.45)	(5.27)
Customer fixed effects	No	No	No	Yes	Yes	Yes
Calendar week fixed effects	No	No	Yes	No	No	Yes
Number of observations	2,629	2,629	2,629	2,629	2,629	2,629

All specifications are done assuming linear probability model

Dependent variable in percentage points

Clustered on customer level standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

6.3 Additional predictions

Note that predictions H4 – H7 are fully descriptive: they simply propose a way we expect different variables to evolve over time. To capture these simple dynamics, univariate regressions suffice, which is why I focus on them in this subsection (Table 7). However, since some of the hypotheses require observing the first purchase of a customer, I restrict the sample to customers whose registration date on the platform is within 2019, which means that we observe their complete history of purchases up to December 31, 2019. There are 1,084 (out of 4,919) such customers.

In addition, to test hypotheses H4 and H5, the measure of riskiness of an item needs to be constructed. To avoid duplicating the return rate variable (which would clearly be largely reflected in the riskiness variable if it was constructed based on the same purchase data), I construct this variable on an apparel category level (t-shirts, dresses, jeans, etc.). Moreover, to calculate riskiness of a category, I use the aggregated data of sales and returns per month for each item — these data were not used in the previous estimation of the return rate. Specifically, I retrieve the item category from its title, group items by category, focus on categories with at least 100 items within them (50 such categories), and mark the categories with the average return rate above median as risky. For example, jeans and dresses are risky according to this measure, while t-shirts and socks are not. The results are robust to other specifications of riskiness (unreported).

The column (1) in Table 7 suggests that the propensity to purchase an item from a risky category increases over time on the sample of higher-end purchases of customers joined in 2019, which is consistent with the individual-level mechanism of uncovering brand quality (H4). The effect however is not robust if estimated on the total sample of customers (unreported). A likely reason of this is that 80 % of higher-end purchases were made by customers who joined the platform earlier (in fact, as Table 2 suggests, about a half of the customers joined before 2017), which means that they have progressed relatively further in discovering the brands. Since a crude category-based riskiness variable is more likely to capture the effect of earlier stages of learning about the brand, the effect is expected to diminish once more experienced customers are added to the sample.

The specification in column (2) refers to all apparel purchases on the subsample of customers for whom we observe the first purchase, and the first apparel purchase belongs to a risky category (as defined above). There are 248 such customers. As the results suggest, the return rate not only does not increase with the number of order for such customers, in fact, it decreases. This is in line with the individual-level model: recall that if the customer starts with the risky item, the return rate in the first period is positive and is equal to zero in the second period, which means that it is decreasing over time, as suggested in H5.

The hypothesis H6 is supported by the column (3): each 10 additional orders move the next order two and a half days closer, which is a large enough effect given that on average a new order is placed each 12.1 days. This effect is robust on the total sample as well — both on total and on apparel purchases. One might argue that such a persistence may be caused by the fact that all customers are becoming more familiar with the platform and in particular the return process, which effectively reduces the hassle cost of a return and incentivizes higher-frequency purchases. However, if that was the case, we would have observed increasing individual return rates for all subsamples, as well as for the entire platform, which we do not (see Tables 3 – 5).

Finally, columns (4) and (5) present evidence in favor of H7. While the effect on the number of items per order is not large (the basket size reduces by one item each 30 orders), it is robust when measured on the total sample. Column (5) suggests that the loyalty may indeed be building over time. In this specification, as opposed to all the previous estimation, the regressor of interest is not the number of order, but the number of weeks since the first purchase (hence reducing the sample to the customers with the observed first purchase). The number of items purchased is aggregated by the number of weeks since the first purchase, and according to the results, it increases over time, so the effect of a lower basket size is offset by the effect of a higher frequency.

Table 7: Tests of additional predictions

	(1)	(2)	(3)	(4)	(5)
	Risky category	*Returned*	*Days from last purchase*	*Basket size*	*Items purchased*
Constant	54.28***	5.65***	15.57***	6.08***	1.21***
	(2.86)	(.57)	(.24)	(.09)	(.01)
Number of order	.348**	-.075**	-.268***	-.033***	
	(.136)	(.026)	(.010)	(.004)	
Number of week					.0030***
					(.0004)
Observations	516	2,483	14,737	15,821	64,479

Specifications (1) and (2) are done assuming linear probability model

Dependent variable in specifications (1) and (2) in percentage points

Clustered on customer level standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

7 Conclusion

In this work, I develop a rational choice theoretical framework that shows how customers make purchase and return decisions in the presence of returns. The model of learning about the return process underscores the role of the hassle cost and captures the population-level dynamics which result in customers with the highest hassle cost leaving the market, so that the average return rate increases. The model of learning about the brand demonstrates that in the presence of returns, customers behave strategically in choosing the order in which to purchase more and less risky categories. This strategic behavior is also reflected in return rates. For brands with high and predictable quality but low and unpredictable fit, this mechanism leads to an increase in individual return rates. The predictions of these models are supported by the data. For lower-end brands, the increase in the return rate is explained by customers' heterogeneity, but for higher-end brands it is not. In addition, the probability of purchasing a risky item increases over time for higher-end brands, and if a customer has purchased a risky item on the first purchase occasion, her return rate decreases. The customers staying in the market (those with lower hassle cost) tend to purchase more frequently, but less items per order. The growing loyalty is reflected in more items purchased per week over time. All these results provide additional evidence in support of validity of the presented models.

One venue for future research based on this framework might be to study its implications for optimal policies of different firms. For instance, if a high-quality firm wants to decrease future return rates, it may need to find ways to offset customers' tendency to buy safer items in the first period. For example, the firm may find it optimal to display riskier items on the home page or in advertisements. This study also provides implications for prediction problems: how to predict the future return rates and purchase dynamics knowing only the information about the first order. As models suggest, the proxies of the customers' hassle cost (such as location or even age), brand type (higher-end versus lower-end), and item attributes (safe versus risky), all provide insights in the dynamics of purchases and returns, which in turn may help firms customize recommendation systems and retention strategies.

A Proof of Lemma 1

Proof. Suppose the realized value of the hassle cost is c_{after}. The customer returns the item if the realized valuation is lower than $-c_{after}$, and otherwise does not return the item and gets v. The expected utility is hence

$$\mathbb{E}U_{after} = -c_{after}\,\mathrm{P}(v < -c_{after}) + \int_{-c_{after}}^{+\infty} x g_v(x)dx \tag{22}$$

Taking $\lim_{c_{after}\to+\infty}$,

$$\begin{aligned}
&\lim_{c_{after}\to+\infty} \mathbb{E}U_{after} = \\
&\lim_{c_{after}\to+\infty} \left(-c_{after}\,\mathrm{P}(v < -c_{after}) + \int_{-c_{after}}^{+\infty} x g_v(x)dx)\right) = \\
&\mathbb{E}v < 0
\end{aligned}$$

Note that if $c_{after} = \bar{c}$,

$$\mathbb{E}U_{after} = -\bar{c}\,\mathrm{P}(v < -\bar{c}) + \int_{-\bar{c}}^{+\infty} x g_v(x)dx > 0 \tag{23}$$

$\mathbb{E}U_{after}$ is strictly decreasing in c_{after}. Let $c_1 < c_2$.
Then

$$\begin{aligned}
&\mathbb{E}U_{after}(c_2) - \mathbb{E}U_{after}(c_1) = \\
&-c_2\,\mathrm{P}(v < -c_2) + c_1\,\mathrm{P}(v < -c_1) + \int_{-c_2}^{+\infty} x g_v(x)dx - \int_{-c_1}^{+\infty} x g_v(x)dx = \\
&(c_1 - c_2)\,\mathrm{P}(v < -c_2) + \int_{-c_2}^{-c_1} (c_1 + x) g_v(x)dx
\end{aligned}$$

Since $x \in (-c_2; -c_1)$, $c_1 + x \in (c_1 - c_2, 0)$, so the second term is negative. The first is also negative, because $c_1 < c_2$. That is, $\mathbb{E}U_{after}(c_2) - \mathbb{E}U_{after}(c_1) < 0$, so $\mathbb{E}U_{after}$ is strictly decreasing in c_{after}.

Since $\mathbb{E}U_{after}$ is also continuous in c_{after}, strictly decreasing and $\mathbb{E}U_{after}(\bar{c}) > 0$, while $\lim_{c_{after}\to+\infty} \mathbb{E}U_{after} < 0$, there exists a unique value of $\tilde{c}$ such that $\mathbb{E}U_{after}(\tilde{c}) = 0$ and $\mathbb{E}U_{after}(c_{after}) > 0 \Leftrightarrow c_{after} < \tilde{c}$.

■

B Proof of Lemma 2

Proof. Suppose the realized value of quality is $\bar{q}$. Then the valuation of the risky item is $v_{risky} = \bar{q} + f$, where f follows g_f. The customer returns the item if the realized valuation is lower than $-c$, that is, $\bar{q} + f < -c$, or $f < -c - \bar{q}$. Otherwise, the customer does not return the item and gets $\bar{q} + f$. The expected utility is hence

$$\mathbb{E}U_2^{risky} = -c\,\mathrm{P}(f < -c - \bar{q}) + \int_{-c-\bar{q}}^{+\infty} (x + \bar{q}) g_f(x) dx \tag{24}$$

Taking $\lim_{\bar{q}\to-\infty}$,

$$\begin{aligned}
&\lim_{\bar{q}\to-\infty} \mathbb{E}U_2^{risky} = \\
&\lim_{\bar{q}\to-\infty} \left(-c\,\mathrm{P}(f < -c - \bar{q}) + \int_{-c-\bar{q}}^{+\infty} x g_f(x) dx + \bar{q}\,\mathrm{P}(f \geq -c - \bar{q}) \right) = \\
0 + &\lim_{\bar{q}\to-\infty} \left(-c\,\mathrm{P}(f < -c - \bar{q}) + \bar{q}\,\mathrm{P}(f \geq -c - \bar{q}) \right) < 0
\end{aligned}$$

since both $-c$ and $\bar{q}$ are negative, while probabilities are positive.

Taking $\lim_{\bar{q}\to+\infty}$,

$$\begin{aligned}
&\lim_{\bar{q}\to+\infty} \mathbb{E}U_2^{risky} = \\
&\lim_{\bar{q}\to+\infty} \left(-c\,\mathrm{P}(f < -c - \bar{q}) + \int_{-c-\bar{q}}^{+\infty} x g_f(x) dx + \bar{q}\,\mathrm{P}(f \geq -c - \bar{q}) \right) = \\
0 + \mathbb{E}f + &\lim_{\bar{q}\to+\infty} \bar{q}\,\mathrm{P}(f \geq -c - \bar{q}) = +\infty
\end{aligned}$$

$\mathbb{E}U_2^{risky}$ is strictly increasing in $\bar{q}$. Let $\bar{q}_1 < \bar{q}_2$.

Then

$$\begin{aligned}
&\mathbb{E}U_2^{risky}(\bar{q}_2) - \mathbb{E}U_2^{risky}(\bar{q}_1) - = \\
&c\,\mathrm{P}(f \in (-c - \bar{q}_2, -c - \bar{q}_1)) + \int_{-c-\bar{q}_2}^{-c-\bar{q}_1} x g_f(x) dx + \bar{q}_2\,\mathrm{P}(f \geq -c - \bar{q}_2) - \bar{q}_1\,\mathrm{P}(f \geq -c - \bar{q}_1) = \\
&\int_{-c-\bar{q}_2}^{-c-\bar{q}_1} (x + c) g_f(x) dx + \bar{q}_2\,\mathrm{P}(f \geq -c - \bar{q}_2) - \bar{q}_1\,\mathrm{P}(f \geq -c - \bar{q}_1) = \\
&\int_{-c-\bar{q}_2}^{-c-\bar{q}_1} (x + c + \bar{q}_2) g_f(x) dx + (\bar{q}_2 - \bar{q}_1)\,\mathrm{P}(f \geq -c - \bar{q}_1)
\end{aligned}$$

Since $x \in (-c - \bar{q}_2; -c - \bar{q}_1)$, $x + c + \bar{q}_2 \in (0, \bar{q}_2 - \bar{q}_1)$, so the first term is positive. The second is also positive, because $\bar{q}_1 < \bar{q}_2$. That is, $\mathbb{E}U_2^{risky}(\bar{q}_2) - \mathbb{E}U_2^{risky}(\bar{q}_1) > 0$, so $\mathbb{E}U_2^{risky}$

is strictly increasing in $\bar{q}$.

Since $\mathbb{E}U_2^{risky}$ is also continuous in $\bar{q}$, strictly increasing and $\lim_{\bar{q}\to-\infty}\mathbb{E}U_2^{risky} < 0$, while $\lim_{\bar{q}\to+\infty}\mathbb{E}U_2^{risky} > 0$, there exists a unique value of $\tilde{q}$ such that $\mathbb{E}U_2^{risky}(\tilde{q}) = 0$ and $\mathbb{E}U_2^{risky}(q) > 0 \Leftrightarrow q > \tilde{q}$.

■

C Additional simulations graphs

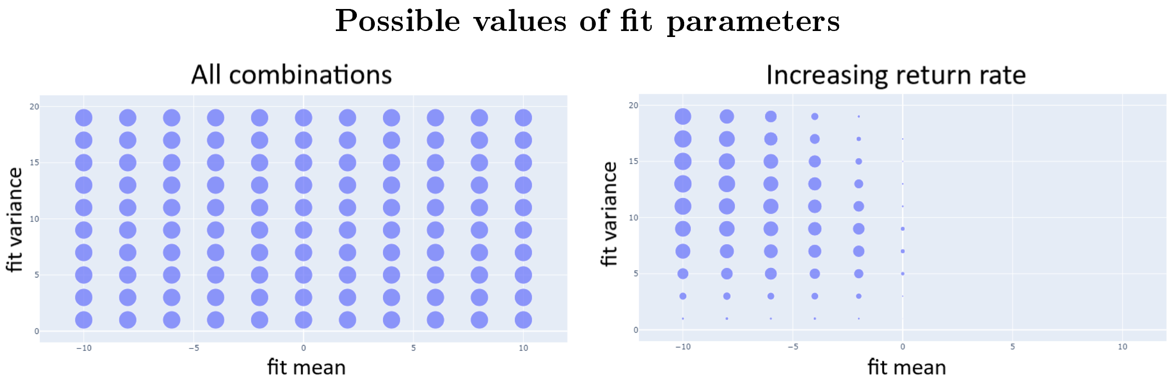

Figure 4: On the left-hand side: all combinations of fit mean and variance that are solved. On the right hand side: combinations of fit mean and variance that result in the increasing pattern in return rate. The marker size corresponds to the number of combinations of these two parameters with the other three that lead to the increasing dynamics.

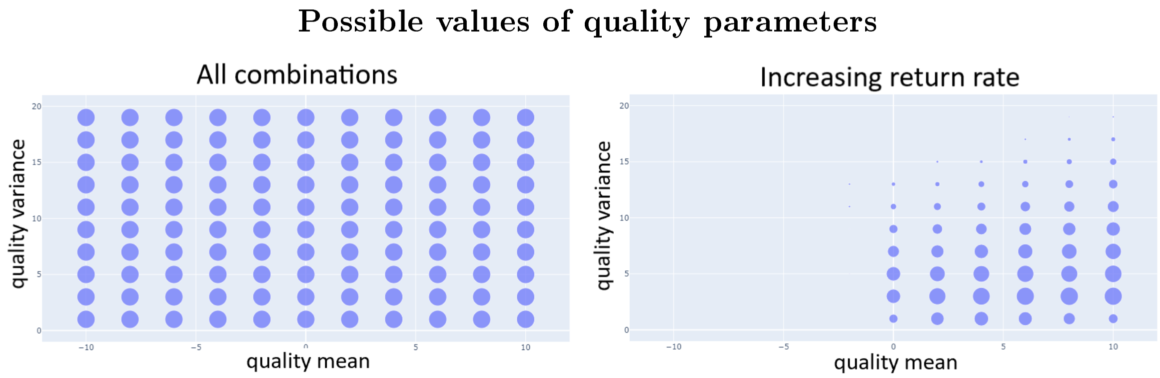

Figure 5: On the left-hand side: all combinations of quality mean and variance that are solved. On the right hand side: combinations of quality mean and variance that result in the increasing pattern in return rate. The marker size corresponds to the number of combinations of these two parameters with the other three that lead to the increasing dynamics.

Possible values of quality variance and return cost

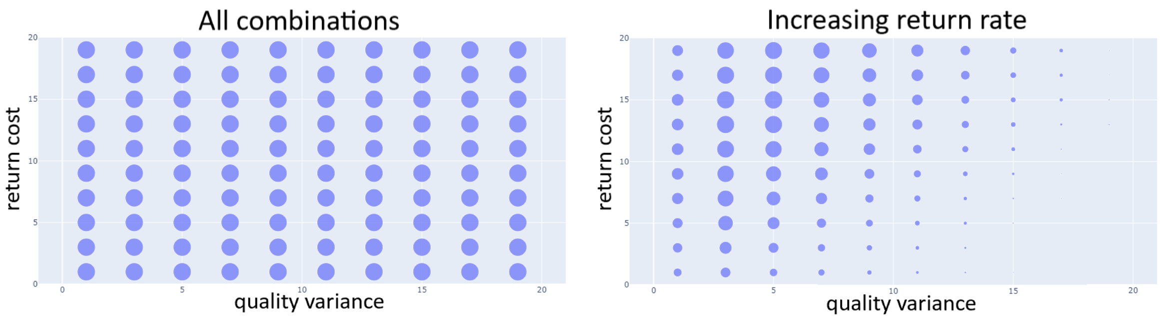

Figure 6: On the left-hand side: all combinations of quality variance and return cost that are solved. On the right hand side: combinations of quality variance and return cost that result in the increasing pattern in return rate. The marker size corresponds to the number of combinations of these two parameters with the other three that lead to the increasing dynamics.

D Additional summary and descriptive tables

Table 8: 10 most common item categories (purchases)

Food	121,295
FMCG nonfood	86,847
Moms & Kids	66,073
Household	63,529
Health & Beauty	62,518
Apparel	62,005
Books	24,519
Zoo	24,519
Pharmacy	22,211
DIY	20,138

Table 9: 10 most common item categories (returns)

Apparel	2,760
Household	705
Moms & Kids	306
DIY	181
Miscellaneous accessories	174
Books	170
Small domestic appliances	153
Sport	148
FMCG nonfood	107
Health & Beauty	104

Table 10: Subsamples of higher- and lower-end brands

High quality brands	Low quality brands
Adidas	oodji
Columbia	Tvoe
Reebok	Concept Club
Puma	Springfield
Lime	Befree
Baon	Sela
Levi's	Women' Secret
United colors of Benetton	Love Republic
U.S. Polo Assn.	
Tom Tailor	
Diesel	
Pepe Jeans	
Tom Farr	
Trussardi Jeans	
adL	
S.Oliver	
Helly Hansen	
Calvin Klein Jeans	

References

Altug, M. S. and T. Aydinliyim (2016). Counteracting strategic purchase deferrals: The impact of online retailers' return policy decisions. *Manufacturing & Service Operations Management, 18*(3), 376-392.

Anderson, E. T., K. Hansen, and D. Simester (2009). The option value of returns: Theory and empirical evidence. *Marketing Science, 28*(3), 405-423.

Arbatskaya, M. (2007). Ordered search. *The RAND Journal of Economics, 38*(1), 119-126.

Bonifield, C., C. Cole, and R. L. Schultz(2010). Product returns on the internet: a case of mixed signals?. *Journal of Business Research, 63*(9-10), 1058-1065.

Branco, F., M. Sun, and J. M. Villas-Boas (2012). Optimal search for product information. *Management Science, 58*(11), 2037-2056.

Burdett, K. and D. A. Malueg (1981). The theory of search for several goods. *Journal of Economic Theory, 24*(3), 362-376.

Carlson, J. A. and R. P. McAfee (1984). Joint search for several goods. *Journal of Economic Theory, 32*(2), 337-345.

Chen, Y. and S. Yao(2017). Sequential search with refinement: Model and application with click-stream data. *Management Science, 63*(12), 4345-4365.

Dzyabura, D., S. El Kihal, J. R. Hauser, and M. Ibragimov (2020). Leveraging the power of images in managing product return rates. Working Paper.

Dzyabura, D. and J. R. Hauser (2019). Recommending products when consumers learn their preference weights. *Marketing Science, 38*(3), 417-441.

Dzyabura, D., S. Jagabathula, and E. Muller (2019). Accounting for discrepancies between online and offline product evaluations. *Marketing Science, 38*(1), 88-106.

El Kihal, S., T. Erdem, and C. Schulze (2021). Customer Return Rate Evolution. Working Paper.

Gatti, J. R. J. (1999). Multi-commodity consumer search. *Journal of Economic Theory*, 86(2), 219-244.

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

Hong, Y. and P. A. Pavlou (2014). Product fit uncertainty in online markets: Nature, effects, and antecedents. *Information Systems Research, 25*(2), 328-344.

Hsiao, L. and Y. J. Chen (2012). Returns policy and quality risk in e-business. *Production and Operations Management, 21*(3), 489-503.

Janakiraman, N., H. A Syrdal, and R. Freling (2016). The effect of return policy leniency on consumer purchase and return decisions: A meta-analytic review. *Journal of Retailing, 92*(2), 226-235.

Jerath, K. (2021). Consumer Search and Firm Return Policy (Don't Be Happy About That Return Policy). Work in Progress.

Kapner, S. (2021a). "Stores Have a Mission: Getting You to Keep That Thing You Bought Online." *Wall Street Journal,* February 9, 2021. www.wsj.com/articles/stores-have-a-mission-getting-you-to-keep-that-thing-you-bought-online-11612866602

Kapner, S. (2021b). "Amazon, Walmart Tell Consumers to Skip Returns of Unwanted Items." *Wall Street Journal,* January 10, 2021. www.wsj.com/articles/amazon-walmart-tell-consumers-to-skip-returns-of-unwanted-items-11610274600

Ke, T. T., Z. J. M. Shen, and J. M. Villas-Boas(2016). Search for information on multiple products. *Management Science, 62*(12), 3576-3603.

Liu, L. and A. Dukes (2013). Consideration set formation with multiproduct firms: The case of within-firm and across-firm evaluation costs. *Management Science, 59*(8), 1871-1886.

Minnema, A., T. H. Bijmolt, J. A. Petersen, and J. D. Shulman (2018). Managing product returns within the customer value framework. In *Customer engagement marketing* (pp. 95-118). Palgrave Macmillan, Cham.

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

Peck, J. and T. L. Childers (2003). To have and to hold: The influence of haptic information on product judgments. *Journal of Marketing, 67*(2), 35-48.

Petersen, J. A. and V. Kumar (2009). Are product returns a necessary evil? Antecedents and consequences. *Journal of Marketing, 73*(3), 35-51.

Shehu, E., D. Papies, and S. A. Neslin (2020). Free Shipping Promotions and Product Returns. *Journal of Marketing Research, 57*(4), 640–658.

Shulman, J. D., A. T. Coughlan, and R. C. Savaskan (2009). Optimal restocking fees and information provision in an integrated demand-supply model of product returns. *Manufacturing & Service Operations Management, 11*(4), 577-594.

Ursu, R. M. (2018). The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science, 37*(4), 530-552.

Ursu, R. M.and D. Dzyabura (2019). Retailers' product location problem with consumer search. *Quantitative Marketing and Economics*, 1-30.

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, 641-654.

Zhou, J. (2014). Multiproduct search and the joint search effect. *American Economic Review, 104*(9), 2918-39.