Journal of Innovation & Knowledge 10 (2025) 100779

Contents lists available at ScienceDirect

Journal of Innovation & Knowledge

journal homepage: www.elsevier.com/locate/jik

Fits like a glove? Knowledge and use of size finders and high-end fashion retail returns

Pankaj C. Patel [*,a], Stefan Karlsson [b], Pejvak Oghazi [c,d,e]

[a] Villanova School of Business, Villanova University, 800 E. Lancaster Ave., Villanova, PA 19085, USA
[b] Faculty of Business, Kristianstad University, Kristianstad, Sweden
[c] School of Social Sciences, Södertörn University, 141 89, Stockholm, Sweden
[d] University of Vaasa, School of Management, PO Box 700, FI-65101, Vaasa, Finland
[e] University of Economics and Human Sciences, Warsaw, Poland

ARTICLE INFO

JEL codes:
D12
L81
M31
C55
C25

Keywords:
retail operations
size
fashion
return supply chain

ABSTRACT

With returns imposing a growing burden on retail supply chains, major e-commerce platforms are increasingly implementing size recommendations to curb returns. Based on a fit valence and fit reference framework, we test whether customers using the size finder are more likely or less likely to return products. We use confidential data from a major fashion e-commerce platform in Sweden that introduced a size finder based on customer-supplied information on weight, build, hips, waist, shoulders, leg-to-torso length, and body shape. In a sample of 496,365 items ordered by 75,707 customers from 113 countries between July 2015 to April 2022, those using the size finder are 0.65% more likely to return an item. The findings are robust to a variety of econometrics tests. Furthermore, machine learning analysis based on gradient boosted trees shows that size finder is among the least important features in predicting returns. However, for each unit quarterly increase in the use of the size finder with purchased items, the customer lifetime value (CLV) increases by 7.51% in the next quarter and 5.53% in the subsequent quarter. Post-hoc interviews with executives in the e-commerce sector demonstrated that, when implementing size recommendation tools, managers in fashion retailers must weigh a small increase in returns against higher CLV from repeat customers.

Introduction

Product returns are an unavoidable, yet costly, feature of online marketplace operations (e.g., Mollenkopf et al., 2011; Abdulla et al., 2019; Ambilkar et al., 2021; Patel et al., 2021; Karlsson et al., 2023). Online returns are projected to rise to about USD 7 billion by 2023 (Ambilkar et al., 2021) with a 2021 National Retail Federation (NRF) study finding that USD 218 billion worth of purchases were returned from U.S. online retail sales of USD 1.05 trillion. In 2020, returns generated 16 million metric tons of CO2 emissions and 5.8 billion pounds of landfill waste.[1] In 2021, the cost of returning a $50 product was roughly $33 including customer care ($2.50), transportation ($5-$7.75), processing ($2.75-$4.00), and liquidation or discounting ($6.25-$35.25).[2] By mid-2022, the escalating bullwhip effect in the post-pandemic era had made the returns problem even more acute.

Yet, returns are a necessary evil for retail operations. On the one hand, returns impose a significant cost on retail operations. According to the Narvar and Return Magic Survey, “41% of customers buy variations of a product with the intent of returning, 42% have returned an online purchase in the last six months, and 89% have returned an online purchase in the last 3 years” (Fuller, 2020). The logistics costs escalate due to increased time, cost, labor, and operational resources (O’Byrne, 2018; Gustafsson et al., 2021). According to some estimates, “managing ‘return and repair’ activities is 10% of total supply chain costs, but if the supply chain overloads, due to inefficient processes, this percentage can grow, reducing profit by 30%.”[3] The personnel at the fashion supply chain’s distribution center do not have sufficient knowledge of reverse logistics, and items are liquidated at deep discounts (Fernie & Grant,

* Corresponding author.
E-mail addresses: pankaj.patel@villanova.edu (P.C. Patel), stefan.karlsson@hkr.se (S. Karlsson), Pejvak.oghazi@sh.se (P. Oghazi).

[1] Source: https://www.optoro.com/2021/02/03/returns-report-powering-resilient-retail-in-2020/
[2] Source: https://www.cbre.com/insights/viewpoints/reverse-logistics-tis-the-stressful-season-for-holiday-gift-returns
[3] Source: https://www.shopify.com/enterprise/ecommerce-returns

https://doi.org/10.1016/j.jik.2025.100779
Received 2 February 2025; Accepted 1 August 2025
Available online 12 August 2025

2019).

On the other hand, according to ReturnLogic, the most profitable customers have approximately an approximate return rate of 32%, indicating that repeat customers who frequently return goods generate more revenues than non-returners. Craig Adkins, the VP of services and operations at Zappos, stated: "Our best customers have the highest return rates, but they are also the ones that spend the most money with us and are our most profitable customers." Customers purchasing the most expensive footwear ultimately return ~50% of everything they buy. [4]

Improving fit perceptions among customers during the purchase phase is important in high-end retail operations. Fit provides a basis for differentiating, and poor fit lowers product value (Ulrich, 2011). We define fit as an ordinal measure where "the levels of the fit attribute can be ordered on a scale" (Wang et al., 2021, p. 268). Return rates are the highest in online fashion apparel with "25-40% overall return rates that can reach up to 75% for specific categories and brands." The primary driver of return is "incorrect fit" followed by non-quantifiable factors such as "it's not me" or "changed my mind" (Nestler et al., 2021, p. 3432). In a survey of 1,000 businesses by ReturnMagic and over 800,000 Shopify customers, preference-based return reasons (e.g., including fit as a preference) tend to drive approximately 72% of all returns in fashion product categories.[5] Fit for clothing refers to size, cut, and shape. For shoes, fit refers to width, support, and flexibility (Fan et al., 2004).

Due to the salience of fit and poor fit as the main reason for returns, the challenge of managing returns in high-end fashion e-commerce platforms is operationally critical. Fashion products are characterized by a shorter product life cycle, increased sensory, status, and perception-based adoption and loyalty, and higher product and emotional dissonance (Brun & Castelli, 2008; Nenni et al., 2013; Caniato et al., 2015). A shorter product life cycle increases inventories, and the frequent launch of newer designs further increases product return costs associated with liquidation and discounting (Stöcker et al., 2021). The processing and transportation of high-end fashion products may require additional expenses. These costs are further exacerbated since customers are more likely to return fashion products, with returns much higher than typical clothing and shoe items (Todeschini et al., 2017; Choi & Guo 2018). In addition to poorer customer experience and higher costs, fashion platforms' size and fit-related returns severely impact the operational supply chain.

To circumvent the return supply chain loop, making sure that customers find the right fit helps to lower returns, thereby reducing the operational burden. During the purchase phase, helping customers find the right fit helps to improve cognitive expectations. Consumers search for products that might meet their needs, forming a cognitive expectation based on reduced inaccuracy in expected sizing (Powers & Jack, 2013; Li & Choudhury, 2021); Miell et al., 2018; Zakaria & Gupta, 2019; Chakraborty et al., 2021). Hong and Pavlou (2014) employed a survey to measure the construct of product fit uncertainty, finding that it exerts stronger positive effects on online product returns than quality uncertainty. Improved size recommendations are salient in improving retail operations, lowering reverse redistribution costs, improving recovery, enhancing customer experience and satisfaction and, most importantly, improving overall supply chain performance (Abdulla et al., 2019; Gustafsson et al., 2021).

Based on the fit valence and fit reference related to consumer expectations theory, we investigate whether the use of an ordinal size finder tool ("where the levels of the fit attribute can be ordered on a scale") on a high-end fashion retailer website reduces returns on a Swedish fashion e-commerce platform (Wang et al., 2021, p. 268). The size finder used by the firm in this study makes sizing recommendations drawn from self-reported information based on weight, build, hips, waist, shoulders, leg-to-torso length, and body shape. In our discussions with the data providers, the reason given for choosing the size finder was to provide a mid-range solution—one that was not as refined as body scanning tools but was as simple as sizing labels indicating "true to fit". Although more advanced technologies are feasible, in our discussion with the executives from the company providing data for this study, it appeared that greater reliance on such technology did not improve fit valence or fit reference. This outcome, in this particular context, was the result of variations in shape and form across designers, and limitations of the technology in discerning flow and form fit in virtual trial rooms. Furthermore, natural language processing and other techniques were highlighted as noisy and, given the faster design turnover, the learning curve was seen as flat. Faster product design turnover adds noise to past learning on preferences and affinities that drive size preferences. The sizing tool in the current study is specific to the industry context and may not be generalizable to standard clothing platforms focused on mass merchandizing.

Based on confidential data on 496,365 items ordered by 75,707 customers from 113 countries between July 2015 and April 2022 at a high-end fashion clothing platform, our findings – controlled for customer fixed effects—show that customers who use the size finder are 0.65% more likely to return an item. To test whether returns vary by customer value quintiles, we compared those in the lowest quintile, who are least likely to return items, with those in the middle quintile, who are more likely to return items. However, those in the upper quintile are not statistically more likely to return a product. The findings are robust when addressing endogeneity through the control function approach, four placebo tests, fixed-effects individual slopes (FEIS), and the post-double selection LASSO. Previous purchases of the same SKU or the same SKU-size combination lowered the odds of a return when using the size finder. Using machine learning, we find that gradient boosted trees have the highest AUC, and the use of the size finder ranks among the least important prediction features. Using customer-quarter aggregated data, for each unit increase in the use of the size finder, the quarterly customer value increases by 7.51% in the next quarter and 5.53% in the subsequent quarter.

A significant body of work has emerged on product returns in operations management (e.g., Abdulla et al., 2019; Ambilkar et al., 2021; Duong et al., 2022). Among operations management scholars and practitioners, interest in managing product returns has grown in recent years (e.g., Abdulla et al., 2019; Ambilkar et al., 2021; Duong et al., 2022). We focus on an early-stage feature of the size finder used in the high-fashion e-commerce platform. Contrary to expectations that customers would prefer to improve fit reference and fit valence, the use of the size finder is less prevalent. The lower effect sizes, using a variety of econometric analyses and machine learning, suggest that providing purchase support through self-reported sizing may produce very limited improvements in returns. Among those who purchase over three successive quarters, the use of the size finder is associated with a meaningful increase in customer purchases. The nominal increase in returns, along with evidence of increased customer value among repeat purchasers, suggests that the gains from the size finder lean more toward improvements in revenue streams, with a slight worsening of operational outcomes from a very small increase in returns.

In the next section, we review the theoretical background on sizing and returns. In Section 3, we present our data and results. In Section 4, we conclude with a discussion of the theoretical and practical implications of our study.

Theoretical background

Research on product returns goes back to the introduction of refunds in the 1960s (Heal, 1980). Refunds were among the early methods of product returns. Over the years, to lower customer purchase risk and increase sales, a variety of product return policies—ranging from money-back guarantees to vouchers and from free returns to lease and purchase agreements—were introduced (Deutsch, 2010). Although

[4] Source: https://www.shopify.com/enterprise/ecommerce-returns

[5] Source: https://www.shopify.com/enterprise/ecommerce-returns

these policies were aimed at lowering risks by reducing the probability of returns or by providing purchase insurance, returns have not only increased exponentially in recent years but have also substantially increased supply chain costs. In the post-pandemic world, product returns have exacerbated the bullwhip effect (Zighan, 2021).

Related to the academic research on returns in operations management, Abdulla et al. (2019) reviewed over 100 papers between 2002 and 2018 on return policies in the context of consumer behavior (Kushwaha, 2015). Braz et al. (2018) and He et al. (2020) reviewed the bullwhip effect from consumer returns. Govindan and Bouzon (2018) and Zailani et al. (2017) appraised the role of stakeholders and factors on reverse logistics driven by product returns. Janakiraman et al. (2016) studied the effect of return policy leniency and return decisions. More recently, Ambilkar et al. (2021) undertook the most comprehensive review of product returns focused on "product recovery, forecasting product returns, consumer behavior, return policy, uncertainty, and technology." Although a full discussion of these rich systematic literature reviews is beyond the scope of this work, the main inferences are: (i) return policies have mixed effects on return outcomes; (ii) logistics, stakeholder, and consumer behavior-related factors add significant challenges to reducing returns; and (iii) newer technologies to manage return forecasts, product designs, and managing purchase decisions are promising but have yet to converge into viable technological ecosystems necessary to lower returns and improve customer experience. Next, we focus on the important element of sizing, a key driver of returns in fashion retail operations.

Sizing in online retail operations

Fit provides a basis for differentiating, and poor fit lowers product value (Ulrich, 2011). We define fit as an ordinal measure with "the levels of the fit attribute can be ordered on a scale" (Wang et al., 2021, p. 268). For clothing, shoes, and accessories, sizing variations are significant across brands, locations, and designs. Considering more well-known brands, the size mapping convention for Reebok is: 6 = 15cm, 7 = 17cm, 8 = 21cm, while for Nike it is: 6 = 16cm, 7 = 18cm, 8 = 22cm (Sembium et al., 2017). Providing size recommendations is not straightforward. In response, three techniques for lowering returns from poor fit have emerged in recent years: (i) structured and unstructured customer feedback text (Isson, 2018); (ii) 2D/3D scanning technology (Daneshmand et al., 2018); and (iii) inter-brand size normalization (Du et al., 2019).[6] All approaches have their merits, but research has shown that each approach also has limitations (Lee & Xu, 2020; Hoque et al., 2021; Lee & Xu, 2022). Using structured and unstructured customer feedback can be less informative because it is based on idiosyncratic customer evaluations. It is subject to bias and can vary depending on contextual conditions such as purchase experience, day, and time of feedback. Similarly, although 2D/3D scanning technology provides a richer fitting experience, the feel, flow, and fit of clothing—especially in high-end fashion retail—may not be reliable (Gill et al., 2022). Although research has shown that 2D/3D technology has promise, it has not been widely adopted by major platforms (Ornati et al., 2022). Finally, in an effort to "standardize sizes", retailers have attempted to make clothing sizes uniform across brands clothing sizes across brands (Du et al., 2019).

Our theoretical foundation is grounded in the fit valence and fit reference literature on consumer behavior (Wang et al., 2021). Fit valence, irrespective of the size label on the item of clothing, is defined in this study as customers' subjective assessment of the fit of a product's attributes based on their personal classification of whether the item is small, well fitting, or oversized (Wang et al., 2021). Fit reference is defined as "body size or dimension" information provided by customers, which describes the circumstances in which they evaluate the product and how they classify their fit preferences—for example, height and weight in the case of clothing. For example, as a fit reference, a customer may identify with a particular body type and have perceptions about certain physiological and personality features. On receiving the item, the customer judges how well it fits using the fit reference; this assessment determines the fit valence. The size finder tool used in the study helps customers assess the item's physical fit and, therefore, enhances their subjective evaluation of fit.

Improving both fit-valence and fit-reference information can help to reduce product-return rates (Wang et al., 2021). Fit valence and fit reference interact in complex ways to influence return behavior. Providing size recommendations based on customer feedback can influence the behavioral aspect of product returns by reducing emotional dissonance. Emotional dissonance is associated with disappointment or sadness resulting from a purchase decision, particularly when considering that acting differently or waiting longer might have delivered a better outcome (Zeelenberg & Pieters, 2004; Alexandra, 2021). Using a size finder to provide details on body structure helps customers reduce emotional dissonance because finding the right fit reduces the disconfirmation gap and lessens the possibility that the purchase will not meet expectations.

Size recommendations are critical on retail fashion platforms (Baier, 2019). Due to higher product turnover from high-frequency design changes, changing customer preferences, and high price points, customers are more likely to be selective in their purchases (Diggins et al., 2016; Stöcker et al., 2021). Fit in high-end fashion is a very important criterion and the most important satisfaction criterion of the shopping experience. Although clothing can be tried on in-store, the difficulty of finding products has led to significant efforts among high-end fashion platforms to recommend sizes. As discussed earlier, fit valence and fit reference (Wang et al., 2021) can be lowered by size recommenders, thus reducing the motivation to return a product. Dissonance and purchase uncertainty are lowered in response (Powers & Jack, 2013; Abdulla et al., 2019). By influencing the pre-choice process between selection and consumption, the size finder may increase disconfirmation between fit valence and fit reference. Customers are helped to avoid unnecessary purchases by increasing expectations of greater dissonance (Bechwati & Siegal, 2005). According to Griffis et al. (2012), size finders can increase later purchases because the customer may experience greater satisfaction from using size finders than from having greater value perceptions and emotional rewards. This factor may increase future product-retention decisions (Simpson et al., 2019).

The challenges to size recommendation efforts are multiplied in the fashion industry where identity, mood, and opinion play an important role (Guigourès et al., 2018). With fashion used to emphasize certain parts of the body, sizing may interact with complex behavioral, individual, and opinion-based factors, adding greater uncertainty to customer satisfaction and leading to higher returns. Due to faster design turnover, fit becomes increasingly important as the introduction of new designs leads to increased deterioration in fit data. Concerns for these variations are likely to be greater in high-end fashion because size concerns may lead customers to buy items in various sizes – hence, lowering customer satisfaction—and fashion retailers to expand their inventories, particularly given the perishability of seasonal fashion products in circumstances of declining patronage. Therefore, size recommendations are important in improving retail operations.

Based on the above discussion, we propose the following research question:

Research question: Does the use of a size finder reduce the probability of returns?

Data and Sample

The data in our research come from a Sweden-based online fashion

[6] Interbrand size normalization refers to synchronizing sizes across brands. For example, a shirt of size 15.5" and 33/34" sleeve collar in brand A is similar in fit to a shirt of size collar 16" and 34/35" sleeve.

clothing company, founded in 2015 and offering clothes mainly for men. The e-tailer aims to transform the fashion industry into a smaller wardrobe with a permanent base collection of 35 items with full environmental transparency. With an extensive offering with more than 15 different size options, the company has aimed for better fit and reduced returns. In 2020, a collection for women was introduced. Germany is its largest market, closely followed by Sweden and the United States. Trousers are the biggest category but have a higher return rate than the company average (45% versus the company average of 37% in Q4 2021).

The data were obtained after signing a confidentiality and non-disclosure agreement. The company shared item-level data for each order and the unique customer identifiers. However, customer demographic details were not shared. We use the fixed-effects specification at the customer level to control for time-invariant characteristics, such as age, gender, and education. We include orders that are completed and exclude exchange orders. We also exclude item categories that do not offer a size finder. These categories are "accessories," "error" (those that cannot be classified), "garment care" and "other". Based on case-wise deletion, our sample includes 496,365 items ordered by 75,707 customers in 113 countries from July 2015 to April 2022.

Description of the size finder

The size finder system developed by the firm seeks to deliver a smoother and engaging customer experience by allowing customers to assess their size for each garment. The firm had 15 different sizes for clothing categories. On clicking "find my size", the buyer is taken through a series of questions on height, body, weight, age, body-to-leg proportion, shoulder, chest shape, belly, preferred fit (tighter, average, looser), and garment length. The size finder is unique to the e-commerce platform and represents an important effort in providing a tailored size finder experience. According to the firm, the size finder has an estimation accuracy of 90% with new (returning) customers having an estimation accuracy of 70% (90%).

Measures

The outcome variable is whether the item is returned (1=yes; 0=no). The company offers free returns on all orders and free shipment for orders over 75 dollars. Customers can return within 30 days and obtain a new product or receive their money back. To manage returns, the company uses a third-party vendor with a digital return registration feature—offering a QR code solution instead of pre-printed return slips—to ease the return process.

Our predictor of interest is whether the size finder, an ordinal fit tool, was used at the time of purchase (0=no; 1=yes). The control variables are the *log of net customer value*, the *quantity of the item ordered*, the *price of the item ordered*, the *weight of the item ordered*, and the *gross margin on the item ordered*. We use the *customer fixed effects* (75,707 customers) to control for a variety of time-invariant customer-related unobservables. We also include item *size* dummies (138 sizes varying by item type), *category* dummies (12 categories), *color* dummies (37 colors varying by item type), *harmonization code* dummies for shipping (31 dummies), *delivery country* dummies (113 countries), *year of the order* dummies (8 years), *week of the year* dummies (52 weeks), the *month of the year* (12 months), and the *quarter of the year dummies* (four quarters). Due to the large number of dummies, we employ a high-dimensional fixed effects model. We use robust standard errors for our analysis.

Descriptives

Table 1 presents the item descriptives. In Table 1(a), of the items ordered, approximately 37.5% were returned, and a size finder was used approximately 7.3% of the time. The average customer value was 391.11 euros, the average price per item was 289 euros, and the gross margin was approximately 78.5%. In Table 1(b), use of the size finder is positively related to item return ($r = 0.045$, $p < 0.01$). Higher customer value is also positively associated with the likelihood of return ($r = 0.044$, $p < 0.01$). Interestingly, the price and quantity of the item ordered were negatively related to the likelihood of return. The higher weight of the item increased the chances of return. Based on the positive correlation between continuous year and whether the item is returned,

Table 1
Sample Descriptives.

Table 1(a). Summary statistics

Definition	Number of items	Mean	SD	Min	p25	Median	p75	Max
The item returned (=1 as yes; =0 as no)	496365	0.375	0.484	0	0	0	1	1
Used size finder	496365	0.073	0.26	0	0	0	0	1
Log of the total net value	496365	5.969	1.417	0	4.998	5.638	6.781	12.85
Quantity of the item ordered	496365	1.076	0.403	1	1	1	1	29
Price of the item ordered	496365	288.967	409.438	0.35	45	95	375	3800
Weight of the item ordered	496365	0.328	0.192	0.06	0.19	0.26	0.47	1.5
Gross margin on the unit ordered	496365	0.785	0.193	-29	0.684	0.723	0.965	1
Year of the order	496365	2020.028	1.418	2015	2019	2020	2021	2022
Week of the year of the order	496365	26.685	14.809	1	14	25	40	52
The month of the year of the order	496365	6.552	3.4	1	4	6	10	12
A quarter of the year of the order	496365	2.51	1.111	1	2	2	4	4

Table 1(b). Pairwise correlations

	Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(9)	(10)	(11)	(12)
1	Item returned (=1 as yes; =0 as no)	1										
2	Used size finder	0.045*	1									
3	Log of total net value	0.044*	-0.065*	1								
4	Quantity of the item ordered	-0.077*	-0.006*	0.071*	1							
5	Price of the item ordered	-0.038*	-0.076*	0.702*	-0.043*	1						
6	Weight of the item ordered	0.144*	-0.060*	0.148*	-0.124*	0.335*	1					
7	Gross margin on the unit ordered	-0.068*	-0.034*	0.505*	0.009*	0.505*	0.019*	1				
8	Year of the order	0.078*	0.238*	0.111*	-0.014*	0.048*	0.161*	-0.007*	1			
9	Week of the year of the order	-0.007*	-0.029*	0.008*	-0.003*	0.023*	0.025*	0.005*	-0.238*	1		
10	Month of the year of the order	-0.007*	-0.029*	0.007*	-0.003*	0.023*	0.024*	0.005*	-0.237*	0.996*	1	
11	Quarter of the year of the order	-0.006*	-0.031*	0.005*	-0.003*	0.020*	0.027*	0	-0.233*	0.967*	0.971*	1

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

returns increased over time ($r = 0.078$, $p < 0.01$). Descriptives by country are available from the authors. Outerwear 15 sizes, knitwear, outerwear 5 sizes, shirts, sweatshirts, and trousers 3 dimensions had among the highest rates of return. In Figure A1(b), The net value by category was generally higher with outerwear items. Gross margins per item were higher or increasing across all categories. Additional details are available from the authors.

Results

The high dimensionality resulting from a large number of fixed effects makes high-dimensional fixed effects more efficient computationally. Therefore, despite the limitations of linear probability models, we use the -reghdfe- routine in Stata 17 for estimates in Table 2 (Correia, 2019). We report the odds ratios and, assuming a 50% chance of returning the item, we compute the probability of returning the item as OR/(1+OR). In models 1 to 3, we present the odds ratios with and without controls. In the first row, we present percentage probability derived from odds ratios presented in the next row. We find that the chances of returning an item increase by 0.65% after using the size finder. Overall, the effects of using the size finder are small.

Robustness checks

Likelihood of returns by customer value

The odds of returning after using the size finder may vary according to customer value. We split the customer value by country x year into quantiles. In models 4 to 8 in Table 2, the chances of return are lower for those in the lower quintile and non-significant for those in the upper quintile. Many of the effects are driven by those in the third and fourth quantiles.

Endogeneity

Endogeneity in the current empirical context stems from omitted variable bias. However, we control for customer value as a proxy for past purchase and return behaviors. Thus, returns may not extemporaneously affect past purchase behaviors. The control function approach, similar to the two-stage least squares model, involves using the exogenous variable—country of origin—to predict the use of the size finder and using the resulting residual from the first step as a control variable in the second step (Wooldridge, 2015).

The countries of origin for the items sold on the platform are Italy, Poland, Portugal, Romania, Tunisia, and Vietnam. These countries represent a meaningful variation in the perceived quality of clothing (Samiee & Chabowski, 2021). A variety of studies have shown that

Table 2
Customer fixed-effects estimates.

	DV = Item returned							
VARIABLES	(1)	(2)	(3)	(4) net customer value quantile by country-year = 1	(5) net customer value quantile by country-year = 2	(6) net customer value quantile by country-year = 3	(7) net customer value quantile by country-year = 4	(8) net customer value quantile by country-year = 5
Percentage probability of return (computed from odds ratio in the next row)	0.65%		0.65%	-0.44%	0.35%	0.56%	0.56%	0.08%
Used size finder (0=no; 1=yes)	1.027***		1.027***	1.023**	1.018**	1.015**	1.024***	0.997
	(0.00339)		(0.00326)	(0.00971)	(0.00804)	(0.00629)	(0.00487)	(0.00446)
Log of net customer value		1.235***	1.235***	1.113***	1.105***	1.108***	1.171***	1.093***
		(0.00184)	(0.00184)	(0.00675)	(0.0145)	(0.0155)	(0.0151)	(0.00458)
Quantity of order		0.961***	0.961***	0.955***	0.960***	0.973***	0.968***	0.989***
		(0.00157)	(0.00157)	(0.00456)	(0.00395)	(0.00373)	(0.00332)	(0.00147)
Price of item		1.000***	1.000***	1.000***	1.000	1.000	1.000	1.000
		(3.29e-06)	(3.29e-06)	(1.57e-05)	(1.01e-05)	(8.15e-06)	(6.68e-06)	(3.50e-06)
Weight of item		1.051***	1.051***	1.110***	0.942**	1.041**	1.066***	1.014
		(0.00941)	(0.00941)	(0.0346)	(0.0237)	(0.0212)	(0.0167)	(0.0107)
Gross margin on the item		0.933***	0.933***	0.976***	0.860***	0.947	0.875***	0.971
		(0.0127)	(0.0128)	(0.00584)	(0.0388)	(0.0322)	(0.0250)	(0.0199)
Customer fixed effects	Included	Included	Included	Included	Included	Included	Included	Included
Size fixed effects	Included	Included	Included	Included	Included	Included	Included	Included
Category fixed effects	Included	Included	Included	Included	Included	Included	Included	Included
Color fixed effects	Included	Included	Included	Included	Included	Included	Included	Included
Harmonization code fixed effects	Included	Included	Included	Included	Included	Included	Included	Included
Delivery country fixed effects	Included	Included	Included	Included	Included	Included	Included	Included
Year of order dummies	Included	Included	Included	Included	Included	Included	Included	Included
Order week of year dummies	Included	Included	Included	Included	Included	Included	Included	Included
Order month of year dummies	Included	Included	Included	Included	Included	Included	Included	Included
Order quarter of year dummies	Included	Included	Included	Included	Included	Included	Included	Included
Standard errors	*robust*	*robust*	*robust*	*robust*	*robust*	*robust*	*robust*	*robust*
Constant	1.452***	0.452***	0.452***	0.748***	0.912	0.837**	0.667***	0.974
	(0.000748)	(0.00579)	(0.00580)	(0.0224)	(0.0663)	(0.0689)	(0.0556)	(0.0333)
Observations	496,365	496,365	496,365	96,859	93,286	96,905	97,472	97,302
R-squared	0.626	0.655	0.656	0.656	0.764	0.816	0.829	0.879
Odds ratios presented								

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

customers take the country of origin into consideration in their purchase decisions (Rashid & Barnes, 2018; Van Esch et al., 2018; Ortega-Egea & García-de-Frutos, 2021) and that they have differing perceptions of clothing made in developing versus developed countries. We anticipate that the country of origin will influence the need to use the size finder. However, this instrument may not directly influence the decision to return the item because the country of origin—for those using it as a purchase criterion—would already be known.

Using the control function approach with the country of origin as an instrument and using quantity, price, weight, size, and category as predictors, Table 3 shows that the effects are higher and more significant compared to the main analysis. This suggests that controlling for residuals after including the country of origin may also influence return behaviors.

Placebo effects

To assess whether the choice of using a size finder is random, we conduct four placebo tests. For the placebo test, we shuffle the use of the finder variable in the following groups: (i) by customer ID; (ii) delivery country; (iii) day of the week; and (iv) month of the year. In Table 4, we find that the effects are significant but negligible for placebo by customer identifier (Model 1, odds ratio: 0.992). The odds ratios for the remaining shuffles were not significant.

Days to return

Various studies state that return customers can be classified as honest returners and renters or those returning items after briefly using them (Altug & Aydinliyim, 2016; Balaram et al., 2022). Renters may be more likely to return later (Balaram et al., 2022). Honest returners will tend to return the item within fewer days because they are more sensitive to receiving the correct size and adding the item to their wardrobe. Hence, they are more likely to return the item sooner.

Using a Tobit model with the coding of 0 days for those not returning—hereby treating non-returned items as censored—we find no support for size finder usage having an effect on the days to return. Full estimates are available from the authors.

Fixed-effects individual and country slopes

As the country and individual preferences may change over time, we use the fixed-effects individual slopes (FEIS) model, which includes time-trend fixed effects at the customer level and the country level (Rüttenauer & Ludwig, 2020). The effects are consistent with the main model. Full estimates are available from the authors.

Table 3
Customer fixed-effects estimates with control function (country of item origin as an instrument).

VARIABLES	(1) DV = Item returned
Used size finder (0=no; 1=yes)	2.991***
	(0.576)
Control function	0.343***
	(0.0662)
Controls	Included
Customer fixed effects	Included
Size fixed effects	Included
Category fixed effects	Included
Color fixed effects	Included
Harmonization code fixed effects	Included
Delivery country fixed effects	Included
Year of order dummies	Included
Order week of year dummies	Included
Order month of year dummies	Included
Order quarter of year dummies	Included
Standard errors	*robust*
Constant	0.416***
	(0.00826)
Observations	496,365
R-squared	0.656
Odds ratios presented	

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

Table 4
Placebo effects (Customer fixed-effects estimates).

VARIABLES	(1) Shuffled by customer	(2) Shuffled by the delivery country	(3) Shuffled by day of the week	(4) Shuffled by month of the year
Used size finder (0=no; 1=yes)	0.992***	1.000	1.001	0.997
	(0.00298)	(0.00189)	(0.00190)	(0.00189)
Log of net customer value	1.235***	1.235***	1.235***	1.235***
	(0.00184)	(0.00184)	(0.00184)	(0.00184)
Quantity of order	0.961***	0.961***	0.961***	0.961***
	(0.00157)	(0.00157)	(0.00157)	(0.00157)
Price of item	1.000***	1.000***	1.000***	1.000***
	(3.29e-06)	(3.29e-06)	(3.29e-06)	(3.29e-06)
Weight of item	1.051***	1.051***	1.051***	1.051***
	(0.00941)	(0.00941)	(0.00941)	(0.00941)
Gross margin on item	0.933***	0.933***	0.933***	0.933***
	(0.0127)	(0.0127)	(0.0127)	(0.0127)
Customer fixed effects	Included	Included	Included	Included
Size fixed effects	Included	Included	Included	Included
Category fixed effects	Included	Included	Included	Included
Color fixed effects	Included	Included	Included	Included
Harmonization code fixed effects	Included	Included	Included	Included
Delivery country fixed effects	Included	Included	Included	Included
Year of order dummies	Included	Included	Included	Included
Order week of year dummies	Included	Included	Included	Included
Order month of year dummies	Included	Included	Included	Included
Order quarter of year dummies	Included	Included	Included	Included
Standard errors	*robust*	*robust*	*robust*	*robust*
Constant	0.453***	0.452***	0.452***	0.452***
	(0.00580)	(0.00579)	(0.00579)	(0.00579)
Observations	496,365	496,365	496,365	496,365
R-squared	0.655	0.655	0.655	0.655

Odds ratios presented
*** p<0.01, ** p<0.05, * p<0.1

Post-double selection LASSO

To lower model mis-specification concerns, we use the post-double selection LASSO (Least Absolute Shrinkage and Selection Operator, Tibshirani, 1996) to estimate parameters in a linear model with many controls. This method combines LASSO and the square-root lasso (Belloni et al., 2014; Ahrens et al., 2019) to improve the robustness of causal inference by controlling for confounding factors.

Highlighting the importance of the post-double selection LASSO, Bonaccolto-Töpfer and Briel (2022) state that the post-double-LASSO estimator identifies controls from a large set of potential regressors—a large number of fixed effects in our models—in a specification with many interactions and polynomials of the returns distribution, thus allowing for more flexible estimation. Using the -pdslasso- routine in Stata 17, the estimates, available from the authors, are consistent with the main inferences.

Does the effect of the size finder vary by item size?

Using the original size categories associated with an item, we test whether returns are conditional on label size. The size finder was not used across all size types. However, starting with model 58, those ordering size L were less likely to return whereas those ordering L-Regular or XS were somewhat more likely to return the item (marginally significant). However, for the remaining size groups in models 61 to 77 and model 79, the effects were positive for most sizes but did not reach statistically significant levels. Overall, we find a small meaningful effect for one size type. However, there is no particular bias in returns after using the size finder for a range of sizes. Full estimates are available from the authors.

The role of past returns behavior

We test for the three types of return behavior: (i) total items returned by the customer before the current order; (ii) the number of times the same SKU was returned before the order; and (iii) the number of times the same SKU-size combination was returned by the customer. The count of past returns had a small effect in the full sample and across customer value quintiles. However, for the number of times the same SKU was returned, using a size finder slightly lowered the odds of return (0.981). This small improvement in returns did not vary across customer value quintiles. Similarly, for the number of times the same SKU size was returned, using a size finder slightly lowered the odds of return (0.972). This small improvement in returns did not vary across customer value quintiles. Overall, the size finder led to small improvements in returns when purchasing and returning the same SKU or the SKU-size combination previously. Full estimates are available from the authors.

Post-hoc analysis—use of the size finder on customer value

Our results show that returns increase by a small amount for those using the size finder. To assess whether the use of the size finder increases customer value, we collapse the data by customer and quarter and use the fixed-effects estimates clustered by customer identifier. Although more quarters are feasible, the data allow us to include only up to two forward quarters due to the lack of customers buying over successive quarters.

In the resulting sample, the use of the size finder in a quarter remains low at 0.288 and a standard deviation of 1.05, with a range from 0 to 48. In interpreting the effects of each unit increase in the use of the size finder, the customer value increases by 7.51% in the next quarter and 5.53% in the subsequent quarter. However, we note that these analyses are based on a limited sample and are meant for exploratory purposes. Full estimates are available from the authors.

Post-hoc analysis—machine learning

Our main analysis shows that the effect of using a size finder is significant but small. We then assess the outcome of returns using machine learning. The purpose of the machine learning approach is to assess whether the use of the size finder is relevant to the model and whether the traditional methods used so far are not biased by Type I errors. This analysis is important because the traditional regression models do not capture the plausible *combinations* of features (variables) that may better explain the return outcomes. We start with the variables employed in the main analysis and use customer dummies.

In the classification model, we use Area Under Curve as a metric to identify the best fit model. Predicting returns is a complex phenomenon driven by a variety of consumer and firm-related factors. In Table 5(a), we used the widely used machine learning models: Naïve Bayes, Generalized Linear Model, Logistic Regression, Fast Large Margin, Deep Learning, Decision Tree, Random Forest, Gradient Boosted Trees, and Support Vector Machines. Although Logistic Regression and generalized linear models have approximately 0.68 AUC, Gradient Boosted Trees have the highest AUC. We use this as our model. In Table 5(b), we present the model estimates for Gradient Boosted Trees, along with the confusion matrix, tree depths, and learning. Related to the relevance of the variables, as expected, the log of the total value was the strongest predictor followed by gross margin, category, and price, as shown in Table 5(c). Consistent with our main findings on small effect size, the use of a size finder was relatively less important—the third least contributing feature. In Table 5(d), we present the decision tree. Here, the log of total net value is the main decision node followed by color (variants) and category of the item. Overall, the machine learning findings further highlight the lower relevance of the use of size finder in predicting returns. The figures are available from the authors on request.

Practitioner reactions to our findings

Our findings demonstrate that size finder has a limited effect on returns but does seem to improve customer lifetime value (CLV). To assess the validity of our findings, we interviewed executives active at various levels of e-commerce operations. Although the practitioners were disappointed to learn about the limited efficacy of the size finder, they valued the indirect benefit of increased CLV. Some of the practitioners did not think the findings were unexpected and felt that, consistent with our theoretical consideration, size finder was not a silver bullet for lowering returns in retail operations. Additional details are available from the second author.

Discussion

The findings paint a complex picture of the advantages derived from efforts to improve ordinal sizing guides as a means to lower returns. Counter-intuitively, the returns increased by a small amount—that is to say, those using a size finder were 0.65% more likely to return items. The findings are robust when addressing endogeneity through the control function approach, four placebo tests, fixed-effects individual slopes (FEIS), and the post-double selection LASSO. Our findings are also supported by the gradient boosted trees approach. Based on the data from customers who regularly made purchases, a unit increase in the size finder is associated with a 5% to 7% increase in net value. The findings can be interpreted as follows. Those using the size finder may have a higher fit requirement and may ascribe greater importance to fit valence. The sensory and visual feedback from trying on fashion items may increase fit valence and make fit reference more salient. “Fit” may be a catch-all reason for non-quantifiable factors driving the decision. When providing details on the size finder, the informational cues on improved fit may increase. This may, in turn, lead to a more intensive evaluation of identity, emotion, and feelings associated with purchase – resulting in a higher percentage of returns by a small amount. The findings on increased net value for customers with consistent purchase behavior could be attributed to the value of personalization provided by the size finder. The size finder is an added attribute for personalization that may increase identity and loyalty associated with the brand (He et al., 2020). Consistent with Gallino and Moreno’s (2018) conclusion that the use of virtual rooms “increases loyalty, helps customers better parse their choice sets, and reduces uncertainty by providing size recommendation” (p. 767), we find that the size finder increases CLV.

Although our inferences are not based on causal estimates, the additional analyses add confidence to the practical value of our estimates. The strength of our analysis, compared to a variety of studies using large-scale data, is that we can control for customer fixed effects, which are robust to fixed-effects individual slopes. Controls for time-invariant customer effects allow for control of personality, gender, and a variety of unobservables. The fixed-effects individual slopes allow for control of trends related to time-invariant effects. Our effect sizes are consistent with broader studies on retail returns (Baumer et al., 2018; Miell et al., 2018; Seyed & Tang, 2019; Forman et al., 2020; Hwangbo et al., 2020; Ogunjimi et al., 2021; Tsagkias et al., 2021).

The bane of retailers in recent years is how to curb the ever-increasing rate of return. Although returns vary by retail industries, product categories, brands, and the extent of omnichannel communication, returns have increased significantly in recent years. Poor size and

Table 5
Machine learning estimates.

Table (a): Model selection.

Model	**AUC**	**Standard Deviation**	**Gains**
Naive Bayes	0.6400	0.0021	-76.0
Generalized Linear Model	0.6759	0.0019	7322.0
Logistic Regression	0.6758	0.0019	7522.0
Fast Large Margin	0.5409	0.0036	-70838.0
Deep Learning	0.5156	0.0058	0.0
Decision Tree	0.5149	0.0001	0.0
Random Forest	0.5753	0.0148	0.0
Gradient Boosted Trees	***0.7060***	***0.0054***	***566.0***
Support Vector Machine	0.5000	0.0000	0.0

Table 5(b): Gradient Boosted Tree model results

MSE	0.1868
RMSE	0.4322
R^2	0.2028
AUC	0.7644
Pr_auc	0.6649
Logloss	0.5526
Mean_per_class_error	0.3150
Default threshold	0.3185

M: Confusion Matrix (Row labels: Actual class; Column labels: Predicted class)

	Item not returned	Item returned	Error	Rate		
Item not returned	20905	14263	0.4056	14,263	/	35,168
Item returned	4730	16352	0.2244	4,730	/	21,082
Totals	25635	30615	0.3377	18,993	/	56,250

Table 5(c): Relevance of variables

Variable	Relative Importance	Scaled Importance	Percentage
Log of total value	6651.1851	1.0000	0.4724
Gross margin on the unit ordered	2051.7632	0.3085	0.1457
Category	1612.4114	0.2424	0.1145
Price of the item ordered	1131.2677	0.1701	0.0804
Customer dummies	752.3081	0.1131	0.0534
Item color	665.3349	0.1000	0.0473
Quantity of the item ordered	367.6890	0.0553	0.0261
Year of the order	220.5787	0.0332	0.0157
Week of the year of the order	213.6858	0.0321	0.0152
A quarter of the year of the order	157.1733	0.0236	0.0112
Weight of the item ordered	112.1401	0.0169	0.0080
Used size finder	110.6946	0.0166	0.0079
The month of the year of the order	25.6089	0.0039	0.0018
Harmonization code	6.4854	0.0010	0.0005

Table 5(d): Tree

```
ln_total_net_value < 5.120
| ln_total_net_value < 4.570
| | variant_name_encode in {Beige,Black Leather,... (14 more)}
| | | variant_name_encode in {Beige,Creme,Light Brown,... (3 more)}: 0.012 {}
| | | variant_name_encode not in {Beige,Creme,Light Brown,... (3 more)}: -0.058 {}
| | variant_name_encode not in {Beige,Black Leather,... (14 more)}
| | | category_encode in {Knitwear,Outerwear 15 Sizes,... (6 more)}: -0.085 {}
| | | category_encode not in {Knitwear,Outerwear 15 Sizes,... (6 more)}: -0.127 {}
| ln_total_net_value ≥ 4.570
| | category_encode in {Belts,Knitwear,Outerwear 5 Sizes,... (5 more)}
| | | customer_id < 57070.500: -0.027 {}
| | | customer_id ≥ 57070.500: 0.006 {}
| | category_encode not in {Belts,Knitwear,Outerwear 5 Sizes,... (5 more)}
| | | customer_id < 14701.500: -0.078 {}
| | | customer_id ≥ 14701.500: -0.042 {}
ln_total_net_value ≥ 5.120
| grossmargin < 0.941
| | ln_total_net_value < 5.694
| | | category_encode in {Trousers 3 Dimensions}: 0.080 {}
| | | category_encode not in {Trousers 3 Dimensions}: 0.009 {}
| | ln_total_net_value ≥ 5.694
| | | ln_total_net_value < 6.123: 0.097 {}
| | | ln_total_net_value ≥ 6.123: 0.159 {}
| grossmargin ≥ 0.941
| | ln_total_net_value < 7.506
| | | ln_total_net_value < 6.800: -0.094 {}
| | | ln_total_net_value ≥ 6.800: -0.054 {}
| | ln_total_net_value ≥ 7.506
```

(*continued on next page*)

Table 5 (*continued*)

Table 5(d): Tree
| | | ln_total_net_value < 10.136: 0.038 {}
| | | ln_total_net_value ≥ 10.136: -0.156 {}

fit are cited as the primary reason for online fashion returns. We drew on fit valence (an overall evaluation of an item's ordinal-fit attribute) and fit reference (perceived ordinal-fit attribute and choice of the product's fit attribute) as the theoretical basis for assessing returns in high-end fashion retailer platforms (Wang et al., 2021). Fit valence and fit reference call for added assessment of high-ticket fashion items, which must be evaluated alongside complementary items (e.g., shoes and accessories). The use of the size finder tool by the e-commerce platform in our data represents an additional attempt to improve personalization and increase the fidelity of fit assessment in the absence of physical examination of size, fit, fabric, design, and pairing.

Contributing to the increasing interest in product returns in operations management (e.g., Abdulla et al., 2019; Ambilkar et al., 2021; Duong et al., 2022), fashion retail platforms pose an added difficulty in appraising purchases, making product returns difficult to forecast because uncertainty in the evaluation of the product intensifies and non-quantifiable elements of purchase evaluation become more salient (Dimoka et al., 2012; He et al., 2020). Moreover, the cost of returns is significantly higher because the remanufacture or refurbishment of high-ticketed fashion items is generally not undertaken. Due to the perishability of designs, the liquidation of items further increases costs. Consumer uncertainty can be resolved somewhat by the size finder. However, its benefits are more discernable in customers of higher value (Kuksov & Lin, 2010).

Managerial implications

Our results call on high-end fashion retailers to cautiously consider investing in size recommendations. The growing pace of product returns, rising return costs, and bullwhip effects from excess inventory call for a closer assessment of how platform features may affect sales, with personalization having the potential to exert a negative financial impact on retail operations. The results highlight the costs and benefits associated with the size finder.

As indicated in Table 2, the percentage chance of returns is 0.65% higher for those using a size finder compared to those not using one. Back of the envelope calculations show that, if 1,000,000 items are sold and 7% use a size finder (as the data shows), then those not using a size finder will return 37% of the time (so, 930,000*.37 = 344,100 returns). Users of the size finder show a return rate of 37.65% above the base rate, which is 0.65% higher. With approximately 70,000 returns from those using the size finder, 0.65 of that figure amounts to 455 items. Based on broader statistics of high return costs—and since there is no system currently to compute return costs—we assume that returns cost about 10% of the average item price (~20 Euros), which amounts to 9,100 euros. This means that the effect is benign, with high-end fashion retailers not losing much by introducing the size finder.

Equally interesting is the increase in the net value of orders by regular users with some purchases over consecutive quarters. Although this analysis is based solely on data from customers purchasing in consecutive quarters, a 5% to 7% increase in net value is meaningful (Table A6). Therefore, in managing cost and benefits tradeoffs, the benefits may be greater for repeat customers, even though return costs may slightly higher.

The findings show that the size recommendation feature minimizes the maximum loss from returns while highlighting the potential upside from repeat customers. Consequently, we call on practitioners to consider both gains and costs from the size recommendation feature. It seems that the size finder has positive implications for increasing sales for repeat customers, although it may not have significant operational cost savings.

Limitations

Our study is not without limitations. First, although our inferences are not causal, the A/B test analysis may not be suitable for making longer-term strategic decisions with operational implications. Our study, by leveraging customer fixed effects over five years, opens a unique window to understanding how size finder tools can impact returns for high-end retailers. A variety of idiosyncratic and contextual factors impact the purchase decisions of the same customers over time, making data collection of A/B tests on the same customers over time less feasible. Although causal estimates using panel data are desirable, firms are generally hesitant to keep control and treated group active over multiple years, and A/B testing windows are generally brief. Focusing on short-term A/B testing approaches provides tactical information in the short term but may not deliver the robust evidence necessary for academic research. We do not discount the value of A/B testing. It is indeed much more valuable than panel data. However, sustained A/B testing for the same research question over time is less feasible in practice. Our discussions with the data provider confirmed this logic.

Second, our study does not focus on the other side of the transaction—the customer. An equally rich body of research on consumer behavior has highlighted the value of customer decision making in returns, and findings remain inconclusive, precisely because of the idiosyncratic nature of returns decisions (Ambilkar et al., 2021; Duong et al., 2022). A variety of demographic and personal factors could provide additional insights into purchase behaviors. Future studies could explore the role of using qualitative studies. We hope that our study offers practical guidance to managers of retail platforms who are considering the deployment of size-fitting recommendations.

Third, the smaller effects and lower variance explained point to the importance of considering the *matter of taste* (Ulrich, 2011; Wang et al., 2021). Although our study does not focus on the complex socio-psychographic or contextual nature of return decisions, the *matter of taste* can be an important driver of such decisions and perhaps may be difficult to predict on the aggregate. The processual elements of returns decisions may be better understood through policy capture studies or through qualitative design approaches that explore narratives on purchase decisions coupled with a temporal bracketing approach. The temporal bracketing approach is particularly helpful in understanding how fit valence and fit reference are used to decompose the purchase process into successive adjacent periods. This allows us to understand how fit valence and fit reference lead to changes in the next stage of the purchase process—especially, product returns.

Fourth, we note that much richer operational analysis is possible. However, in our discussions with the executives at the company and those who provided quotations interpreting our results, the item- and order-level costs are difficult to estimate, unlike a manufacturing setup where costs are more easily traceable. The challenges relate to complex combinations of the various service levels required to manage returns, the perishability of fashion clothing items, and the vendors available to liquidate the items. These factors make it difficult to estimate detailed operational cost breakdowns. A related point concerns the inventory of units. In our conversation with executives, inventory storage and insurance costs in the fashion industry are not significant due to the small volume of units. Given that the units are produced in smaller batches—with more frequent seasonal batches—concerns about significant inventory costs are limited. We reiterate that the smaller change in returns shows that operational costs from returns do not change.

Nevertheless, this does not discount the value of focusing further on retail operations in the fashion industry.

In conclusion, the "lack of proper fit" is the main reason for returns in fashion retail. Variations in cognitive, contextual, and idiosyncratic preferences drive fit valence and fit reference, which increase product returns. This study is based on fine-grained data from a major fashion retailer that introduced a size finder. This is a mid-range option that does not increase expectations associated with scanning technologies, which are not as simple as sizing charts and not as reliant on the sizing recommendations of natural language processing. Our findings paint a complex picture—although the use of a size finder increases returns by a small amount, CLV is increased for repeat customers.

CRediT authorship contribution statement

Pankaj C. Patel: Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Stefan Karlsson:** Writing – review & editing, Resources, Data curation. **Pejvak Oghazi:** Writing – review & editing.

References

Ahrens, A., Hansen, C. B., & Schaffer, M. (2019). *PDSLASSO: Stata module for post-selection and post-regularization OLS or IV estimation and inference* [Computer software]. *Boston College Department of Economics*.

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

Ambilkar, P., Dohale, V., Gunasekaran, A., & Bilolikar, V. (2021). Product returns management: A comprehensive review and future research agenda. *International Journal of Production Research*, 1–25.

Baier, S. (2019). Analyzing customer feedback for product fit prediction (arXiv preprint arXiv:1908.10896). arXiv.

Balaram, A., Perdikaki, O., & Galbreth, M. R. (2022). Bracketing of purchases to manage size uncertainty: Should online retailers be worried? *Naval Research Logistics, 69*(1), 3–22.

Baumer, E. P. S., Sun, R., & Schaedler, P. (2018). Departing and returning: Sense of agency as an organizing concept for understanding social media non/use transitions. In *Proceedings of the ACM on Human-Computer Interaction, 2(CSCW)* (pp. 1–19).

Bechwati, N. N., & Siegal, W. S. (2005). The impact of the prechoice process on product returns. *Journal of Marketing Research, 42*(3), 358–367.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies, 81*(2), 608–650.

Bonaccolto-Töpfer, M., & Briel, S. (2022). The gender pay gap revisited: Does machine learning offer new insights? *Labour Economics, 77*, Article 102223.

Braz, A. C., De Mello, A. M., de Vasconcelos Gomes, L. A., & de Souza Nascimento, P. T. (2018). The bullwhip effect in closed-loop supply chains: A systematic literature review. *Journal of Cleaner Production, 202*, 376–389.

Brun, A., & Castelli, C. (2008). Supply chain strategy in the fashion industry: Developing a portfolio model depending on product, retail channel and brand. *International Journal of Production Economics, 116*(2), 169–181.

Caniato, F., Crippa, L., Pero, M., Sianesi, A., & Spina, G. (2015). Internationalisation and outsourcing of operations and product development in the fashion industry. *Production Planning & Control, 26*(9), 706–722.

Chakraborty, S., Hoque, M. S., Rahman, J. N., Biswas, M. C., Bardhan, D., & Lobaton, E. (2021). Fashion recommendation systems, models and methods: A review. *Informatics, 8*(3), 49.

Choi, T.-M., & Guo, S. (2018). Responsive supply in fashion mass customisation systems with consumer returns. *International Journal of Production Research, 56*(10), 3409–3422.

Correia, S. (2019). *REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects [Computer software]*. Boston College Department of Economics.

Daneshmand, M., Helmi, A., Avots, E., Noroozi, F., Alisinanoglu, F., Arslan, H. S., Gorbova, J., Haamer, R. E., Ozcinar, C., & Anbarjafari, G. (2018). 3D scanning: A comprehensive survey (arXiv preprint arXiv:1801.08863). arXiv.

Deutsch, T. (2010). Exploring new insights into retail history. *Journal of Historical Research in Marketing, 2*(1), 130–138.

Diggins, M. A., Chen, C., & Chen, J. (2016). A review: Customer returns in fashion retailing. In T. M. Choi (Ed.), *Analytical modeling research in fashion business* (pp. 31–48). Springer.

Dimoka, A., Hong, Y., & Pavlou, P. A. (2012). On product uncertainty in online markets: Theory and evidence. *MIS Quarterly, 36*(2), 395–426.

Du, E. S. J., Liu, C., & Wayne, D. H. (2019). Automated fashion size normalization (arXiv preprint arXiv:1908.09980). arXiv.

Duong, Q. H., Zhou, L., Meng, M., Van Nguyen, T., Ieromonachou, P., & Nguyen, D. T. (2022). Understanding product returns: A systematic literature review using machine learning and bibliometric analysis. *International Journal of Production Economics, 243*, Article 108340.

Fan, J., Yu, W., & Hunter, L. (2004). *Clothing appearance and fit: Science and technology*. Woodhead Publishing.

Fernie, J., & Grant, D. B. (2019). *Fashion logistics: Insights into the fashion retail supply chain*. Kogan Page Publishers.

Forman, J., Dogan, M. D., Forsythe, H., & Ishii, H. (2020). DefeXtiles: 3D printing quasi-woven fabric via under-extrusion. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20)* (pp. 1222–1233). Association for Computing Machinery.

Fuller, D. (2020). *Ingredients for ecommerce success*. Aquitude Press.

Gallino, S., & Moreno, A. (2018). The value of fit information in online retail: Evidence from a randomized field experiment. *Manufacturing & Service Operations Management, 20*(4), 767–787.

Gill, S., Januszkiewicz, M., & Ahmed, M. (2022). Digital fashion technology: A review of online fit and sizing. In N. Zakaria (Ed.), *Digital manufacturing technology for sustainable anthropometric apparel* (pp. 135–163). Woodhead Publishing.

Govindan, K., & Bouzon, M. (2018). From a literature review to a multi-perspective framework for reverse logistics barriers and drivers. *Journal of Cleaner Production, 187*, 318–337.

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

Guigourès, R., Ho, Y. K., Koriagin, E., Sheikh, A.-S., Bergmann, U., & Shirvany, R. (2018). A hierarchical Bayesian model for size recommendation in fashion. In *Proceedings of the 12th ACM Conference on Recommender Systems* (pp. 392–396). ACM.

Gustafsson, E., Jonsson, P., & Holmström, J. (2021). Reducing retail supply chain costs of product returns using digital product fitting. *International Journal of Physical Distribution & Logistics Management, 51*(8), 877–896.

He, Y., Xu, Q., & Wu, P. (2020). Omnichannel retail operations with refurbished consumer returns. *International Journal of Production Research, 58*(1), 271–290.

Heal, G. (1980). Spatial structure in the retail trade: A study in product differentiation with increasing returns. *The Bell Journal of Economics, 11*(2), 565–583.

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

Hoque, M., Rasiah, R., Furuoka, F., & Kumar, S. (2021). Technology adoption in the apparel industry: Insight from literature review and research directions. *Research Journal of Textile and Apparel, 25*, 292–307.

Hwangbo, H., Kim, E. H., Lee, S.-H., & Jang, Y. J. (2020). Effects of 3D virtual "try-on" on online sales and customers' purchasing experiences. *IEEE Access, 8*, 189479–189489.

Isson, J. P. (2018). *Unstructured data analytics: How to improve customer acquisition, customer retention, and fraud detection and prevention*. John Wiley & Sons.

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

Karlsson, S., Oghazi, P., Hellström, D., Patel, P. C., Papadopoulou, C., & Hjort, K. (2023). Retail returns management strategy: An alignment perspective. *Journal of Innovation & Knowledge, 8*(4), Article 100420.

Kuksov, D., & Lin, Y. (2010). Information provision in a vertically differentiated competitive marketplace. *Marketing Science, 29*(1), 122–138.

Kushwaha, S. (2015). Consumers and product returns in reverse supply chain: A literature review. *IUP Journal of Supply Chain Management, 12*(4), 36.

Lee, H., & Xu, Y. (2020). Classification of virtual fitting room technologies in the fashion industry: From the perspective of consumer experience. *International Journal of Fashion Design, Technology and Education, 13*(1), 1–10.

Lee, H., & Xu, Y. (2022). Influence of motivational orientations on consumers' adoption of virtual fitting rooms (VFRs): Moderating effects of fashion leadership and technology visibility. *International Journal of Fashion Design, Technology and Education*, 1–11.

Li, M., & Choudhury, A. H. (2021). Using website information to reduce postpurchase dissonance: A mediated moderating role of perceived risk. *Psychology & Marketing, 38*(1), 56–69.

Miell, S., Gill, S., & Vazquez, D. (2018). Enabling the digital fashion consumer through fit and sizing technology. *Journal of Global Fashion Marketing, 9*(1), 9–23.

Mollenkopf, D. A., Frankel, R., & Russo, I. (2011). Creating value through returns management: Exploring the marketing–operations interface. *Journal of Operations Management, 29*(5), 391–403.

Nenni, M. E., Giustiniano, L., & Pirolo, L. (2013). Demand forecasting in the fashion industry: A review. *International Journal of Engineering Business Management, 5*, 37.

Nestler, A., Karessli, N., Hajjar, K., Weffer, R., & Shirvany, R. (2021). SizeFlags: Reducing size and fit related returns in fashion e-commerce. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* (pp. 3432–3440). ACM.

O'Byrne, R. (2018). Reverse logistics—Who does it well? *MHD Supply Chain Solutions, 48*(4), 30–31.

Ogunjimi, A., Rahman, M., Islam, N., & Hasan, R. (2021). Smart mirror fashion technology for the retail chain transformation. *Technological Forecasting and Social Change, 173*, Article 121118.

Ornati, M., Picco-Schwendener, A., & Marazza, S. (2022). Sizing up the body: Virtual fit platforms in fashion e-commerce. *International Journal of Fashion Studies, 9*(1), 199–218.

Ortega-Egea, J. M., & García-de-Frutos, N. (2021). Mapping the influence of country-of-origin knowledge, consumer ethnocentrism, and perceived risk on consumer action against foreign products. *Journal of Consumer Behaviour, 20*(5), 1164–1178.

Patel, P. C., Baldauf, C., Karlsson, S., & Oghazi, P. (2021). The impact of free returns on online purchase behavior: Evidence from an intervention at an online retailer. *Journal of Operations Management, 67*(4), 511–555.

Powers, T. L., & Jack, E. P. (2013). The influence of cognitive dissonance on retail product returns. *Psychology & Marketing, 30*(8), 724–735.

Rashid, A., & Barnes, L. (2018). Country of origin association in retail and wholesale branding. *Journal of Promotion Management, 24*(3), 312–331.

Rüttenauer, T., & Ludwig, V. (2020). Fixed effects individual slopes: Accounting and testing for heterogeneous effects in panel data or other multilevel models. *Sociological Methods & Research, 52*(1), 43–84.

Samiee, S., & Chabowski, B. R. (2021). Knowledge structure in product-and brand origin–related research. *Journal of the Academy of Marketing Science, 49*(5), 947–968.

Seyed, T., & Tang, A. (2019). Mannequette: Understanding and enabling collaboration and creativity on avant-garde fashion-tech runways. In *Proceedings of the ACM on Designing Interactive Systems Conference (DIS '19)* (pp. 1–13). ACM.

Simpson, D., Power, D., Riach, K., & Tsarenko, Y. (2019). Consumer motivation for product disposal and its role in acquiring products for reuse. *Journal of Operations Management, 65*(7), 612–635.

Stöcker, B., Baier, D., & Brand, B. M. (2021). New insights in online fashion retail returns from a customers' perspective and their dynamics. *Journal of Business Economics, 91* (8), 1149–1187.

Todeschini, B. V., Cortimiglia, M. N., Callegaro-de-Menezes, D., & Ghezzi, A. (2017). Innovative and sustainable business models in the fashion industry: Entrepreneurial drivers, opportunities, and challenges. *Business Horizons, 60*(6), 759–770.

Ulrich, K. T. (2011). *Design: Creation of artifacts in* society (SSRN Scholarly Paper No. 1951106). *Social Science Research Network*.

Van Esch, P., Northey, G., Duffy, S., Heller, J., & Striluk, M. (2018). The moderating influence of country of origin information seeking on homophily and product satisfaction. *Journal of Promotion Management, 24*(3), 332–348.

Wang, Y., Ramachandran, V., & Liu, S. O. R. (2021). Do fit opinions matter? The impact of fit context on online product returns. *Information Systems Research, 32*(1), 268–289.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources, 50*(2), 420–445.

Zailani, S., Govindan, K., Shaharudin, M. R., & Kuan, E. E. L. (2017). Barriers to product return management in automotive manufacturing firms in Malaysia. *Journal of Cleaner Production, 141*, 22–40.

Zakaria, N., & Gupta, D. (2019). *Anthropometry, apparel sizing and design*. Woodhead Publishing.

Zeelenberg, M., & Pieters, R. (2004). Beyond valence in customer dissatisfaction: A review and new findings on behavioral responses to regret and disappointment in failed services. *Journal of Business Research, 57*(4), 445–455.