Marketing Science

The Option Value of Returns: Theory and Empirical Evidence

Eric T. Anderson, Karsten Hansen, Duncan Simester,

Please scroll down for article—it is on subsequent pages

informs.

The Option Value of Returns: Theory and Empirical Evidence

Eric T. Anderson, Karsten Hansen
Marketing Department, Kellogg School of Management, Northwestern University, Evanston, Illinois 60208
{eric-anderson@kellogg.northwestern.edu, karsten-hansen@kellogg.northwestern.edu}

Duncan Simester
MIT Sloan School of Management, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02142, simester@mit.edu

When a firm allows the return of previously purchased merchandise, it provides customers with an option that has measurable value. Whereas the option to return merchandise leads to an increase in gross revenue, it also creates additional costs. Selecting an optimal return policy requires balancing both demand and cost implications. In this paper, we develop a structural model of a consumer's decision to purchase and return an item that nests extant choice models as a special case. The model enables a firm to both measure the value to consumers of the return option and balance the costs and benefits of different return policies.

We apply the model to a sample of data provided by a mail-order catalog company. We find considerable variation in the value of returns across customers and categories. When the option value is large, there are large increases in demand. For example, the option to return women's footwear is worth an average of more than $15 per purchase to customers and increases average purchase rates by more than 50%. We illustrate how the model can be used by a retailer to optimize his return policies across categories and customers.

1. Introduction

The high incidence of product return costs manufacturers and retailers dearly, with an estimated $100 billion lost annually through product depreciation and management of the returns process (Blanchard 2005). Yet the incidence of returns is at least partially attributable to firms' own policies. Retailers could reduce the incidence by making returns more difficult or more costly for customers. Managers recognize this opportunity in interviews but express concerns that implementing more costly return policies may upset their customers and reduce demand. Surprisingly, none of the retailers that we interviewed had varied their return policies and measured this trade-off.

A possible explanation for the lack of measurement is that return policies are set by the market, with every retailer forced to adopt the same policy. However, a review of return policies across different retailers reveals wide variation in their policies, even within the same product market. For example, while apparel retailer Coldwater Creek allows returns of any merchandise at any time, but many of its competitors require that returns occur within 30 or 60 days.

The variation in policies is even starker when it occurs within the same firm. There are numerous examples of a single firm implementing different policies in different product categories: Sears has more restrictive return policies for home electronics and mattresses than for other categories (Merrick and Brat 2005). There are even examples of policies varying across customers: Macy's elite and platinum cardholders do not have to pay the cost of return postage in contrast to other Macy's customers who do incur this cost.[1]

Selecting an optimal return policy requires balancing both demand and cost implications. A liberal return policy that allows customers to return unwanted merchandise increases customers' expected

---

[1] At an informal level, discussion with colleagues revealed many anecdotes in which retailers apparently relaxed their return policies for frequent customers. There is also at least one documented example of a retailer banning customers because their propensity to return items was too high. In 2003, *The Boston Globe* reported that two sisters had been banned from all 21 stores in the Filene's Basement department store chain because of "a history of excessive returns" (Mohl 2003). Similarly, many catalog retailers suppress catalog mailings to customers who return a high proportion of their items.

utility. Such a policy increases gross demand but also leads to more returns and greater costs. A restrictive return policy reduces the value of this option to customers, reducing both returns and demand. Although retailers are often aware of this trade-off, measuring the resulting costs and benefits is difficult. Whereas most remote retailers have developed policies for varying catalog content through randomized split sample tests, they generally lack the infrastructure to vary post-transaction return policies across different randomly selected customer samples. In this paper, we offer an alternative solution to the problem. We develop a structural model of customer demand and returns that nests extant models as a special case. The model is estimated on historical transaction data and we then simulate the impact of varying a retailer's return policy on customer behavior and derive the resulting profit implications. The profit calculation explicitly trades off the opportunity cost of lost demand when returns are more difficult with the additional product depreciation and administrative costs when the return policy is relaxed.

A key generalizable contribution of this paper is that we provide a tool that allows managers to identify opportunities to vary return policies across product categories and customer segments. An important output of the model is that for each customer we quantify the option value of returning unwanted merchandise. This option value varies across customers due to differences in demand and both intrinsic and extrinsic return costs. The option value also varies across product categories, reflecting different levels of customer uncertainty about product fit together with differences in customer preferences. In an empirical application, using data from a mail-order apparel catalog, we confirm that these option values are large and can have a significant impact on demand. For example, in the women's footwear category, the opportunity to return unwanted footwear is valued at more than $15 per transaction, which is 30% of the average price and increases net category demand by more than 50%. More restrictive return policies would hurt both demand and profits in this category. Yet, in other categories such as men's tops, customers place less value on the return option ($3), so that additional restrictions on returns could increase firm profits. Our model also distinguishes customers who place little value on the return option from those customers who place a lot of value on the option. Retailers may use these estimates to help evaluate customer-specific, category-specific, or channel-specific return policies. For example, a retailers may want to relax their return policies to customers who purchase over the Internet while tightening their policies for customers who purchase in their stores.

We also compare the demand elasticities derived from our model against two benchmark models. Previous demand models generally either ignore returns and focus on *gross demand* or they net out returns and estimate *net demand*. Estimates from our structural model suggest that demand is approximately 10% more elastic compared to a model of net demand. Further investigation reveals that differences in elasticity estimates can be as high as 30% and depend in part upon how much customers value the return option. Importantly, the direction of these differences is not systematic and can be either positive or negative. We show that these differences can be explained in part by how our demand model accounts for the return option.

An important aspect of the structural model is that we estimate the return cost for each consumer, which leads to two important contributions. First, estimates of this parameter allow a deeper understanding of customer behavior. Second, the measures facilitate targeted marketing policies. For example, suppose a customer purchases many items but never returns any of them. Is this a customer with very high return cost? Or is this a customer who has a very strong preference for a retailer's products? Parameter estimates from our structural model separate these two explanations, which allows a manager to respond with appropriate marketing tactics. If the observed behavior is due to high return costs, a firm may respond by subsidizing these costs. These subsidies would be wasted on a customer who has a strong intrinsic preference for all of the firm's products.

1.1. Previous Literature

In this paper, we focus on customer returns for reasons of taste and fit. This focus is distinct from retailer to manufacturer returns (Pasternack 1985, Padmanabhan and Png 1997, Emmons and Gilbert 1998) and remanufacturing returns (Rogers and Tibben-Lembke 2001). The process of managing the flow of returned items has received considerable attention in the inventory management literature (Kiesmüller and van der Laan 2001, Savaskan et al. 2004) but is beyond the scope of this paper. Finally, whereas customer return fraud is a concern for many retailers (Walker 2004), we assume that all customer returns are legitimate.

A series of theoretical papers have previously investigated the interaction between retail return policies and customer taste and fit. A premise of these models is that customers purchase an item with incomplete information that is later resolved via postpurchase inspection. Several papers identify a common trade-off on margin versus volume that retailers face (Davis et al. 1995, 1998; Che 1996). Return policies reduce customer risk, which allows retailers to raise prices, but a customer will return an item if

the price exceeds her ex post valuation, which will reduce demand. Che (1996) and Shulman et al. (2009) show that retail information about product fit can serve the same role as postpurchase inspection. As a result, customers may have a higher willingness to pay in the absence of product information than with additional information. A related finding can also be found in Heiman et al. (2001) who show that product demonstrations can potentially reduce retailer profits. Finally, Che (1996) shows that a return option facilitates risk sharing, which may increase social welfare if customers are risk averse.

Empirical research on retail return policies and customer return behavior has been more limited. Davis et al. (1995) show that fashion retailers are more likely to accept returns of "regularly priced" merchandise than "clearance" items and interpret this as evidence that return policies are more liberal when the product has a higher salvage value. Davis et al. (1998) analyze the return policies of 133 retailers. They show that retailer return policies vary with how quickly a product is consumed, the salvage value of returned merchandise, and whether there are opportunities to cross-sell or substitute other items when returns occur.

Hess and Mayhew (1997) is one of the few empirical studies that measures individual customer return behavior. Unlike our model, the authors do not consider customer purchase decisions and instead focus on predicting whether and when a return occurs. Using both actual and simulated data, they show that a split adjusted hazard model is better at predicting return behavior than a regression model. A limitation of Hess and Mayhew's (1997) model is that they do not allow for unobserved customer heterogeneity, which limits how well their model can predict individual customer return behavior.

More recently, Anderson et al. (2009) identify two factors that may influence customer return behavior when an item is sold at a low price. First, they show that lower prices lead to additional consumer surplus (i.e., perceived value), which reduces the likelihood that a customer will return an item. Second, they argue that low prices may attract customers with different return propensities (i.e., customer heterogeneity). In turn, changes in the mix of customers may affect the number of returns. They find empirical support for these predictions, which validates a key assumption in our structural model that consumers are less likely to return a lower-priced item. Moreover, we explicitly allow for the types of heterogeneity in customer return rates that Anderson et al. (2009) document.

A separate though related stream of literature has interpreted retailers' return policies as a signaling mechanism. For example, Moorthy and Srinivasan

(1995) show that a liberal return policy can act as a credible signal of product quality. These models require that the retailer has private information about product quality; otherwise, the retailer has no information to signal. As a result, the unobserved product features tend to be "quality" rather than "fit" characteristics, and so these findings are perhaps more closely related to the literature on warranty returns than returns for reasons of taste and fit. More generally, the signaling literature raises the possibility that a return policy can affect a customer's quality and price perceptions.

This paper also contributes to an emerging research stream that recognizes the need to coordinate marketing and operations decisions (Ho and Tang 2004, Karmarkar 1996). Customer return policies affect both customer demand and operational costs and, hence, require tremendous intrafirm coordination. While there is a recognized need by academics, empirical research on the issue remains scarce. Notable exceptions include Kulp et al. (2004), who conduct a large-scale survey to investigate the value of interfirm coordination between manufacturers and retailers, and Anderson et al. (2006), who document the long-run costs of stockouts.

The remainder of this paper is organized as follows. In §2, we present a structural model of customer return behavior and describe a strategy for identification. We introduce an empirical application of the model in §3 and present findings in §4. The implications are discussed in §5, and the paper concludes in §6.

2. A Model of Purchase Incidence and Returns

In this section, we develop an econometric model of purchase incidence and return behavior. The model derives the option value of returns at the customer level. This in turn makes it possible to predict how demand will respond to changes in a retailer's return policies. As we show later in this section, the model generalizes to a standard purchase incidence model when returns are banned by a firm. Before describing the model mathematically, we offer an overview of the customer and firm behavior captured in the model.

We develop this model in the context of a single firm selling to many customers, but the model can be extended to incorporate competition. A derivation of the competitive model is available in the appendix.

2.1. Model Overview
The model interprets customer returns as evidence of poor fit between the product and a customer's preferences. Consider the following example: A customer is

considering the purchase of a Havasu blue bath towel from a mail-order catalog. Prior to purchase, the customer can inspect a picture of the product and read a brief product description. When the towel is delivered, the customer learns the exact color, texture, size, and design. In addition, the customer can now ascertain whether the towel matches other bath items. For some customers, the Havasu blue towel is a fantastic fit, but for others it is a poor fit, in which case some customers may return the item.

The key element of this example is that fit is not fully observed by the customer prior to purchase. This assumption is well suited to remote retail settings where customers select a product from a catalog or Internet site without being able to physically inspect the product. The assumption may also hold in a more traditional retail setting if physical inspection within the store is not sufficient to guarantee suitability. Examples might include home furnishings that need to match existing furnishings in the home, apparel that needs to match other apparel in the wardrobe, and gifts and other items where the fit extends to someone other than the purchaser. Moreover, some consumers may prefer to try on an apparel item for proper fit at home rather than in a store. Examples include women who shop with their children or customers who are uncomfortable trying on merchandise in public spaces.

The requirement that fit is not fully known prior to purchase also suggests that the assumption is more applicable to durable goods markets, in which it is rare for customers to purchase the same product on more than one occasion. If the customer had already purchased the Havasu blue towel, there would be little uncertainty about product fit prior to the subsequent purchase. Of course, fit is generally specific to a product so that previously purchasing the Havasu blue towel does not reveal the fit of the Bermudian blue sheets.

In the model, we consider a two-stage process. In stage 1, customers decide whether to order an item and then in stage 2, after receiving the item they decide whether to keep or return it. To develop the model, we focus our exposition on a single product category but the generalization to multiple categories is straightforward. To simplify notation, we do not include a product subscript and readers should remember that customers are purchasing a single product from a large category of items.

2.2. Model
Consider customer i who is deciding in period t whether to return or keep an item. We assume that

$$U(\text{return})_{it} = -R_i, \tag{1}$$

$$U(\text{keep})_{it} = \mu_{it} + \psi_{it} + \varepsilon_{it}, \tag{2}$$

where

$$\mu_{it} \equiv \beta_i' X_t. \tag{3}$$

The utility of returning an item is simply the return cost and this is always negative (i.e., $U(\text{return})_{it} < 0$).[2] We assume that the return cost varies across customers but does not vary over time. Our definition of a customer's return cost includes both the monetary cost of a return such as shipping or mailing fees, and the psychic return costs such as hassle and time investment. The model specification does not explicitly distinguish between different types of return costs, but more detailed customer data may facilitate such decomposition. The utility of keeping an item has three components. The first term μ_{it} is the deterministic utility that is known by both the researcher and consumer at the time of purchase. We assume that the vector X_t contains marketing variables that impact the mean utility level such as price and promotional information. The vector X_t also contains a constant, which allows for a customer-specific intercept that does not vary over time. The third term ε_{it} is a standard econometric error term that is known to the customer prior to purchase but is not observed by the researcher. This error term captures time-varying shocks to preferences.

The focus of the model is the second term in Equation (2), ψ_{it}, which measures the fit of the transaction. For apparel products, this could be the physical fit of the product or it could be sensory related such as the color or texture of a fabric. We assume that ψ_{it} is never observed by the researcher and is only observed by the customer after receiving the product. Motivated by our empirical application, we interpret ψ_{it} as a fixed product characteristic that is time-invariant. The parameter ψ_{it} has a time subscript because in our application the available products vary each period: short-sleeved shirts are sold in summer and long-sleeved shirts are sold in fall. The reader should keep in mind that other interpretations of ψ_{it} are possible. For example, in other applications, damage to a product in shipping may be a concern and ψ_{it} may capture this time-varying product characteristic. Alternative interpretations of ψ_{it} do not affect our model development but may influence model estimation.[3]

Whether a consumer keeps an item depends on the net utility compared to returning the item. We refer to this as U_{it}^K, which is defined as

$$U_{it}^K = \mu_{it} + \psi_{it} + R_i + \varepsilon_{it}. \tag{4}$$

---

[2] If a consumer can exchange an item, then the return of an unwanted item may facilitate the purchase of a desired item. In this case, the net utility of a return may be positive. Our model does not consider exchanges or this type of learning.

[3] In our empirical application, we consider purchase incidence from a category of items (e.g., women's tops). In other applications where there are unobserved, time-varying product effects (e.g., damage in shipping), one might consider purchase of an individual item.

The product is kept if $U_{it}^K > 0$. For the consumer, this is a deterministic decision because μ_{it}, ψ_{it}, R_i, and ε_{it} are all known after product inspection.

We now turn to the first stage to derive the probability that a customer places an order. Prior to purchase, ψ_{it} is unknown but all customers know that this random variable has the following distribution:

$$\psi_{it} \sim N(0, \sigma_\psi^2). \tag{5}$$

Consistent with our earlier description of consumer behavior, we assume that the variance of ψ_{it} is constant over time. This assumption precludes customers from learning about unobserved taste and fit over time. The subscript i on the random variable should remind the reader that each customer receives his own realization of the random variable each period. Thus, while a firm may sell the same yellow shirt to different consumers, each person will receive his own independent realization of ψ_{it}.

A consumer has the option of returning an item that has poor fit and incorporates this in her purchase decision. We can write the expected utility from purchasing an item as

$$E[U_{it}(\text{order})] = E[U(\text{keep})_{it} \mid \text{keep}] \Pr(\text{keep})$$
$$+ E[U(\text{return})_{it} \mid \text{return}] \Pr(\text{return}), \tag{6}$$

where expectations are with respect to ψ_{it}. Recall that from the consumer's perspective, ψ_{it} is the only source of uncertainty in the purchase decision. Using the properties of the truncated normal distribution, we can show that the first term on the right-hand side of (6) is equal to

$$E[U(\text{keep})_{it} \mid U_{it}^K > 0]$$
$$= \mu_{it} + \varepsilon_{it} + E[\psi_{it} \mid \psi_{it} > -(\mu_{it} + \varepsilon_{it} + R_i)]$$
$$= \mu_{it} + \varepsilon_{it} + \sigma_\psi \frac{\phi((\mu_{it} + \varepsilon_{it} + R_i)/\sigma_\psi)}{\Phi((\mu_{it} + \varepsilon_{it} + R_i)/\sigma_\psi)}. \tag{7}$$

Customer i expects to keep the order with probability

$$\Pr(\text{keep}) = \Pr(U_{it}^K > 0) = \Pr(\psi_{it} > -(\mu_{it} + \varepsilon_{it} + R_i))$$
$$= \Phi((\mu_{it} + \varepsilon_{it} + R_i/\sigma_\psi). \tag{8}$$

The expected utility of returning an item can be written as

$$E[U(\text{return})_{it} \mid \text{return}] = E[-R_i \mid U_{it}^K < 0] = -R_i. \tag{9}$$

Finally, the probability of returning an item is simply one minus the probability of keeping an item. Substituting these expressions into (6) yields

$$E[U_{it}(\text{order})] = \Phi\left(\frac{\mu_{it} + \varepsilon_{it} + R_i}{\sigma_\psi}\right)[\mu_{it} + \varepsilon_{it} + R_i]$$
$$+ \sigma_\psi \phi\left(\frac{\mu_{it} + \varepsilon_{it} + R_i}{\sigma_\psi}\right) - R_i$$
$$= H(\mu_{it} + \varepsilon_{it}, R_i, \sigma_\psi). \tag{10}$$

The above expression characterizes the consumer's purchase decision. The researcher does not observe ε_{it} and we need to integrate over this to obtain the probability of observing customer i ordering at time t. This is given by

$$\Pr(E[U_{it}(\text{order})] > 0 \mid \mu_{it}, R_i, \sigma_\psi)$$
$$= \Pr(H(\mu_{it} + \varepsilon_{it}, R_i, \sigma_\psi) > 0)$$
$$= \int 1_{\{H(\mu_{it} + \varepsilon_{it}, R_i, \sigma_\psi) > 0\}} \phi(\varepsilon_{it} \mid 0, \sigma_\varepsilon^2) \, d\varepsilon_{it}. \tag{11}$$

At first glance, the expression for the order probability in Equation (11) does not appear very tractable. However, by analyzing the properties of the H function, we can derive an alternative expression for this probability that achieves two goals. First, the alternative expression allows us to compare our model with the standard purchase incidence model and, second, the simplification facilitates estimation. We begin by defining H as a function of $\omega = \mu + \varepsilon$ and σ_ψ:

$$H(\omega, R, \sigma_\psi) = \Phi\left(\frac{\omega + R}{\sigma_\psi}\right) \times [\omega + R]$$
$$+ \sigma_\psi \phi\left(\frac{\omega + R}{\sigma_\psi}\right) - R. \tag{12}$$

To simplify notation, we omit the consumer i and time t subscripts from Equation (12). One can easily show that the H function has a negative lower bound, a positive upper bound, and increases monotonically with ω. This implies that H as a function of ω has a unique root (i.e., it equals zero for only one value of ω), and this root can be quickly located numerically by a Newton algorithm. Define this root as $\varpi(R, \sigma_\psi)$:

$$H(\varpi(R, \sigma_\psi), R, \sigma_\psi) = 0. \tag{13}$$

Note that $\varpi(R, \sigma_\psi)$ is the value of $\mu + \varepsilon$, which makes the customer exactly indifferent between ordering and not ordering. H is monotonic in ω, which in turn implies that

$$\omega > \varpi(R, \sigma_\psi) \iff H(\omega, R, \sigma_\psi) > 0,$$
$$\omega < \varpi(R, \sigma_\psi) \iff H(\omega, R, \sigma_\psi) < 0. \tag{14}$$

An order occurs iff $\omega = \mu + \varepsilon > \varpi(R, \sigma_\psi)$. Therefore, we define

$$U_{it}^O = -\varpi(R_i, \sigma_\psi) + \mu_{it} + \varepsilon_{it}^O, \tag{15}$$

where $U_{it}^O > 0$ if an order is placed and $U_{it}^O < 0$ otherwise. We can now write the probability of an order as

$$\Pr(U_{it}^O > 0 \mid R_i, \mu_{it}, \sigma_\psi) = \Phi\left(\frac{-\varpi(R_i, \sigma_\psi) + \mu_{it}}{\sigma_\varepsilon}\right). \tag{16}$$

Equation (16) is a lot simpler than (11) and easily lends itself to a comparison with a standard purchase incidence model.[4] Finally, we let $\varepsilon_{it}^O = \varepsilon_{it}$ and $\varepsilon_{it}^K = \varepsilon_{it} + \psi_{it}$ so that

$$\varepsilon_{it}^O, \varepsilon_{it}^K \mid \Sigma \sim N(0, \Sigma) \qquad (17)$$

and

$$\Sigma = \begin{pmatrix} \sigma_\varepsilon^2 & \sigma_\varepsilon^2 \\ \sigma_\varepsilon^2 & \sigma_\varepsilon^2 + \sigma_\psi^2 \end{pmatrix}. \qquad (18)$$

For each customer, we have a β_i vector and one return cost parameter R_i. We assume that

$$\begin{pmatrix} \beta_i \\ \log(R_i) \end{pmatrix} \sim N\left(\begin{matrix} \bar{\beta} \\ \bar{R} \end{matrix}, \Omega \right). \qquad (19)$$

This specification imposes the restriction that return costs cannot be negative. This is an important restriction to impose because otherwise the H function defined in Equation (10) will not have a root. We next compare this joint model of demand and returns to the standard model of purchase incidence.

2.3. Comparison to the Standard Purchase Incidence Model

The traditional purchase incidence model does not consider the possibility of returns and instead assumes that the latent utility of a purchase is simply

$$U_{it} = \mu_{it} + \varepsilon_{it}. \qquad (20)$$

The resulting purchase probability is

$$\Pr(U_{it}^O > 0 \mid \mu_{it}, \sigma_\varepsilon) = \Phi\left(\frac{\mu_{it}}{\sigma_\varepsilon}\right), \qquad (21)$$

where σ_ε is customarily set to one to aid identification. Comparing this to the order probability in the joint model (16), we see that there is an added term $-\varpi$ in the intercept. Recall that we earlier defined this term as minus the unique root of (12). It satisfies the following properties:

$$-\varpi(R, \sigma_\psi) \geq 0, \qquad (22)$$

$$\lim_{R \to 0} -\varpi(R, \sigma_\psi) = \infty, \quad \lim_{R \to \infty} -\varpi(R, \sigma_\psi) = 0, \qquad (23)$$

$$\lim_{\sigma_\psi \to 0} -\varpi(R, \sigma_\psi) = 0, \quad \lim_{\sigma_\psi \to \infty} -\varpi(R, \sigma_\psi) = \infty. \qquad (24)$$

As $R \to \infty$ or $\sigma_\psi \to 0$, the joint model converges to the traditional model. In other words, as returns get increasingly expensive or as the uncertainty about product fit declines, the two models become identical. When returns are too expensive, zero consumers will exercise the return option. Similarly, when there

is no uncertainty about product fit, there will never be returns: all products shipped to a customer are kept. We also observe that the order rate increases as customer return costs decrease or as the uncertainty of product fit increases. In the limit as $R \to 0$ or $\sigma_\psi \to \infty$, customers purchase in every time period. However, in both situations, returns also increase.

A more precise characterization of these effects can be given by the following expressions for the probabilities of the three possible outcomes:

1. *No Order*:

$$\Pr(U_{it}^0 < 0 \mid R_i, \mu_{it}) = 1 - \Pr(U_{it}^0 > 0 \mid R_i, \mu_{it}, \sigma_\psi)$$

$$= 1 - \Phi\left(\frac{-\varpi(R_i, \sigma_\psi) + \mu_{it}}{\sigma_\varepsilon}\right). \quad (25)$$

2. *Order and Return*:

$$\Pr(U_{it}^0 > 0, U_{it}^K < 0 \mid R_i, \mu_{it})$$

$$= \int_{(-\varpi(R_i, \sigma_\psi) + \mu_{it})/\sigma_\varepsilon}^{\infty} \Phi\left(-\frac{(R_i + \mu_{it} + \sigma_\varepsilon \varepsilon)}{\sigma_\psi}\right) \phi(\varepsilon) \, d\varepsilon. \quad (26)$$

3. *Order and Keep*:

$$\Pr(U_{it}^O > 0, U_{it}^K > 0 \mid R_i, \mu_{it})$$

$$= \int_{(-\varpi(R_i, \sigma_\psi) + \mu_{it})/\sigma_\varepsilon}^{\infty} \Phi\left(\frac{(R_i + \mu_{it} + \sigma_\varepsilon \varepsilon)}{\sigma_\psi}\right) \phi(\varepsilon) \, d\varepsilon. \quad (27)$$

It is now straightforward to derive the limits on these probabilities when R or σ_ψ go to zero or infinity. These limits further illustrate the relationship between our model and the standard purchase incidence model:

$$\lim_{R \to \infty} \Pr(U_{it}^O < 0 \mid R_i, \mu_{it})$$

$$= \lim_{\sigma_\psi \to 0} \Pr(U_{it}^O < 0 \mid R_i, \mu_{it}) = 1 - \Phi\left(\frac{\mu_{it}}{\sigma_\varepsilon}\right), \quad (28)$$

$$\lim_{R \to 0} \Pr(U_{it}^O > 0, U_{it}^K < 0 \mid R_i, \mu_{it})$$

$$= \lim_{\sigma_\psi \to 0} \Pr(U_{it}^O > 0, U_{it}^K < 0 \mid R_i, \mu_{it}) = 0. \quad (29)$$

Thus, in the "costly return" scenario or the "no uncertainty" scenario, order rates converge to those of the standard purchase incidence model and return rates go to zero. Similarly,

$$\lim_{R \to 0} \Pr(U_{it}^O < 0 \mid R_i, \mu_{it})$$

$$= \lim_{\sigma_\psi \to \infty} \Pr(U_{it}^O < 0 \mid R_i, \mu_{it}) = 0, \quad (30)$$

$$\lim_{R \to 0} \Pr(U_{it}^O > 0, U_{it}^K < 0 \mid R_i, \mu_{it})$$

$$= \int_{-\infty}^{\infty} \Phi\left(-\frac{[\mu_{it} + \sigma_\varepsilon \varepsilon]}{\sigma_\psi}\right) \phi(\varepsilon) \, d\varepsilon, \quad (31)$$

$$\lim_{\sigma_\psi \to \infty} \Pr(U_{it}^O > 0, U_{it}^K < 0 \mid R_i, \mu_{it}) = \frac{1}{2}. \quad (32)$$

Notice that when consumer return costs are zero, the return rates are different than in the infinite uncertainty scenario. However, it is still the case that varying R and σ_ψ in opposite directions has a similar impact on the likelihood function.[5]

2.4. The Option Value of Returns

The option to return mismatched merchandise always provides positive value to the customer. From the firm's perspective, a return policy provides value when it leads to an increase in net demand or profits. Although the return option always increases orders, it may not increase net demand (orders minus returns) if the increase in returns outweighs the increase in orders. Next, we develop metrics of the customer's option value and its impact on net demand.

To find a customer's option value, we need to compare ex ante expected utility when returns are banned, i.e., $R = \infty$, to ex ante utility under the finite R case. We first consider the case where returns are banned and customers keep all purchased items. Once an item arrives and is inspected, the ex post utility equals $\omega + \psi$. Because ψ has mean zero and all items are kept when $R = \infty$, the ex ante expected utility of ordering equals ω.

When R is finite, the ex ante expected utility equals H, which is defined in Equation (12). Note that $H > \omega$ so that customers always place a positive value on the return option. Define π as the decrease in customer utility such that a customer is indifferent between buying when returns are allowed and returns are banned. The value of π is given by the solution to the following equation:

$$H(\omega - \pi, R, \sigma_\psi) = \omega. \tag{33}$$

Thus, π is the amount of utility a customer receives when a firm offers an option of returning unwanted merchandise. Although it may not be immediately apparent because we have omitted subscripts, it is important to recognize that this option value varies across customers and over time (i.e., π_{it}). The customer variation follows from heterogeneity in preferences. For example, the disutility of returns R_i varies by customer and reduces the option value. The temporal variation follows from the econometric error term ε_{it}, which is known to the customer ex ante. For example, if ε_{it} is very large, a customer is unlikely to return an item and the option value is low. Let $\bar{\pi}_i = \mathrm{E}[\pi_{it}]$, which is the average option value (in utils)

for customer i. To convert this to a dollar metric, we divide by a customer's price sensitivity.

From the firm's perspective, the value of offering returns rests on whether net demand increases. If a return policy increases orders but many items are returned, it may not be profitable to offer a return option. We define $\Delta(\mu, R)$ as the change in net demand when returns are offered:

$$
\begin{aligned}
\Delta(\mu, R) &\equiv \Pr(U^O > U^K > 0 \mid \mu, R) \\
&\quad - \lim_{R \to \infty} \Pr(U^O > 0, U^K > 0 \mid \mu, R) \\
&= \Pr(U^O > 0, U^K > 0 \mid \mu, R) \\
&\quad - \Pr(U^O > 0 \mid \mu).
\end{aligned}
\tag{34}
$$

Similar to π, $\Delta(\mu, R)$ will vary across customers and over time. When $\Delta(\mu, R)$ is aggregated over all customers over a specific time horizon, it captures the change in total firm demand when returns are offered. In our paper, we focus on the change in monthly net demand, but the metric could readily be extended to a longer time horizon.

The impact on net demand can be negative. To illustrate this possibility, consider a base case where returns are banned and contrast this with a scenario where returns are allowed. In the base case, all purchased items are kept so gross demand equals net demand. When returns are allowed, customer orders increase because $\pi > 0$, but at the same time there are more returns: the base case has zero returns. For net demand to increase, the number of incremental orders must exceed the number of returns. When this fails to hold, then net demand decreases and offering returns has negative value to the firm.

2.5. Identification

To begin our discussion of identification, we make the standard assumption that $\sigma_\varepsilon = 1$. In the typical discrete-choice model, this assumption is usually sufficient for identification. However, our model of customer returns introduces two additional parameters R and σ_ψ. Unlike the standard purchase model, a customer's return decision also provides us with additional information. The critical question is whether this additional information is sufficient for identifying these additional parameters.

Unfortunately, for a single category, the additional information provided by customer returns is not sufficient for identification. To illustrate this point, we first recognize that one property of $\varpi(R, \sigma_\psi)$ is homogeneity of degree one: $\lambda \varpi(R, \sigma_\psi) = \varpi(\lambda R, \lambda \sigma_\psi)$. This implies that the likelihood is invariant to scale transformations, so that $L(\theta_h, R_h, \sigma_\varepsilon) = L(\lambda \theta, \lambda R_h, \lambda \sigma_\psi, \lambda \sigma_\varepsilon)$ for any $\lambda > 0$. Further, our analysis of the limits of R and σ_ψ on the outcome probabilities in Equations (28)–(32) illustrates that these parameters

---

[5] Readers may wonder why the return rate converges to $\frac{1}{2}$ in the infinite variance case. In this case, the distribution of ψ converges to a distribution with infinite mass in the tails, which implies that the probability of any finite interval is zero. By assumption, the distribution of σ_ψ is symmetric around zero, and this leads to exactly half of all orders being returned.

have similar impacts on the likelihood. Thus, we cannot separately identify these parameters. A careful numerical study confirms that the likelihood function plateaus in the (R, σ_ψ) dimension. We conclude that without additional information our model is not identified for a single category.

To solve this identification problem, we pursue a multicategory approach in which we jointly estimate the model for several product categories. We then make the assumption that a consumer's return cost is invariant across these categories. Formally, we assume that

$$R_{ic} = R_i, \quad c = 1, \ldots, C, \tag{35}$$

where R_{ic} is the return cost of customer i for category c. We believe this assumption is appropriate in our empirical application of customer return behavior for apparel items but that it may not generalize to other settings. We provide support for this assumption in our empirical application.

Finally, we also assume that

$$\sigma_{\psi, c'} = 1 \tag{36}$$

for one category c'. This allows us to compare the relative uncertainty of product fit across categories. To see why these assumptions lead to identification, consider category c'. The restriction $\sigma_{\psi, c'} = 1$ yields an estimate of the return cost R_i for each customer. By assuming that return costs do not vary across categories, we can then estimate product uncertainty for the remaining product categories.

One limitation of this identification strategy is that estimation requires that we have at least some customers who purchase in more than one category. In our empirical application, we limit our attention to only multicategory buyers. However, including a mix of single-category buyers and multicategory buyers would be sufficient.

We augment the specification of the unobservable heterogeneity distribution to account for multiple product categories. For each customer, we have a β_{ic} vector for each category and one return cost parameter R_i. We collect all these parameters in the vector $\theta_{ic} = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{iC}, \log R_i)$ and specify the distribution of θ_{ic} as

$$\begin{pmatrix} \beta_{i1} \\ \vdots \\ \beta_{ic} \\ \log(R_i) \end{pmatrix} \sim N\left(\begin{pmatrix} \bar{\beta}_1 \\ \vdots \\ \bar{\beta}_c \\ \bar{R} \end{pmatrix}, \Omega \right) = N(\Pi Z, \Omega). \tag{37}$$

Because we place no restrictions on the covariance matrix Ω, this specification allows preferences and marketing mix sensitivities to be correlated across categories. In addition, return costs may be correlated

with preferences and marketing mix sensitivities in all categories. In the empirical application, we obtain estimates of both the population parameters $(\bar{\beta}_1, \bar{\beta}_2, \ldots, \bar{\beta}_c, \bar{R})$ and the customer level parameters θ_{ic}. The regression structure ΠZ allows us to investigate whether the population parameters vary systematically with customer characteristics and category characteristics. If we include category characteristics in Z, we must still impose the identifying restriction that the return rate is constant across categories.

3. Empirical Application

To illustrate the value of the information that the model provides, we consider an application to a mail-order apparel retailer. The company sells primarily private label men's and women's clothing at moderate price points, and most products carry the firm's own brand name. Products are sold exclusively through company-owned distribution channels, which include catalogs, an Internet site, and retail stores. The company maintains a strict policy of charging the same prices across all of its channels.

For the purposes of this illustration, we focus on a sample of 987 customers who purchase from three broad categories of items: women's tops, men's tops, and women's footwear. The database includes a total of 1,087 customers, but we use a random sample of 100 customers as a holdout sample. We selected these three categories as demand tends to be independent across the categories, so that a purchase in one category does not affect demand in another category. Selecting categories in this manner avoids the added complexity of accounting for cross-category demand effects. Generalizing the model to allow for cross-category demand effects is an area of future research.

For apparel items, the selling season is typically 10–15 weeks in duration and new items are regularly introduced into a category. For example, the summer season features men's short-sleeved shirts and the winter season features men's long-sleeved shirts. Across years, there is variation in patterns, material, and designs for all three categories. This regular introduction of new products ensures that there is uncertainty about product fit prior to purchase.

The company has a relatively liberal return policy. Purchases from any channel can be returned at any of the company's stores including its factory outlets. Customers can also return by mail using a prepaid preaddressed label that they can download off the company's Internet site. Customers using this option are charged a small fee to cover the cost of the return postage. There are no time limits on how long a customer can hold an item before returning it and no requirement that the product be defective. Customers

Table 1   Summary of Customers' Transactions

	Women's tops	Men's tops	Women's footwear
Average probability of an order (in a month)	0.196	0.147	0.122
Return rate: prob(Return | Order)	0.233	0.140	0.291
Average probability of order and keep (in a month)	0.150	0.126	0.087
Average number of orders per year	2.35	1.76	1.46
Average interpurchase period (months)	4.48	5.96	6.84
Number of customers	987	987	987

are also not required to contact a customer service representative in advance to authorize the return and they are not charged any restocking fees. Conversations with managers confirm that there were no major changes in the retailer's return policies during the period of our study.

We observe customer purchases over a period of 126 months ($10\frac{1}{2}$ years). Although we observe transactions on a daily basis, customers typically place at most a single order in any 30-day period. Thus, we use one month as the time interval for our analysis.[6] Each customer's first purchase from the company occurs on or after week 1 in our sample and so we observe each customer's entire purchase history (eliminating any left-censoring issues). Although the company operates retail stores, the customers we consider make few purchases through this channel over the period of our data. To simplify the model, we assume that a consumer's information about product fit does not vary across channels. We hope to relax this assumption in later work.

We summarize customer purchase behavior in the three categories in Table 1. The average probability that a customer orders in any given month is 0.196, 0.147, and 0.122 for women's tops, men's tops, and women's footwear, respectively. Importantly, customer return rates differ dramatically across these three categories: men's tops have the lowest return rate (14%) while women's footwear have the highest return rate (29%). Returns reduce the conversion rate, which is the probability that an ordered item is kept. We will later use these probabilities as a benchmark by comparing how they change as we vary return costs. The average interpurchase time ranges from 4.48 months for women's tops to 6.84 months for women's footwear.

Recall that our empirical identification strategy assumes that return costs R_i are independent across

---

[6] We model purchase and return incidence. If a customer purchases more than one item from a category in a month, we treat this as a single purchase. If a customer returns more than one item in a category, we treat this as a return.

categories. An alternative assumption is that return costs are concave, decreasing in the number of items returned. For example, if a customer is returning shoes, it may be lower cost to return a shirt. Estimation of a concave return cost function at the consumer level requires variation in return behavior for each consumer. In our sample, over 50% of the consumers never return two or more items in a single period, and so the estimation of a concave return cost function is not feasible for a substantial fraction of customers.

Under our assumption, return rates should be invariant to whether a customer purchased a single item or multiple items. Under the alternative assumption, return rates should be greater when a customer purchases multiple items. In a separate analysis, which is available from the authors, we found no significant relationship between return rates and the number of items purchased. This suggests that our independent return cost assumption is plausible in this application. We caution that the result was not unexpected as we chose three categories that were expected to yield independent decisions; purchase and return decisions for men's tops and women's tops are not expected to be related. However, our independence assumption may need to be relaxed in other applications.

Our specification of the price variable is guided by previous research, which shows that consumers may be sensitive to both the absolute price level and the percent discount from the regular price (Kahneman and Tversky 1984, Lichtenstein and Bearden 1989, Alba et al. 1994). This has led researchers to include variables for both the absolute price level and the percent discount (i.e., price promotion) in demand models (Jedidi et al. 1999). In our application, we defined the regular price variable as the average regular price of items sold in a category each month. Similarly, the percent discount variable was operationalized as the average percent discount of all items sold in a category each month. We computed these variables on a database of transactions that included a large sample of the retailer's customers.

In preliminary analysis, we included both price variables in our model specification. Consistent with past research, we found that customers were more sensitive to the percent discount than to the regular price (Blattberg et al. 1995). On average, the regular price sensitivity was near zero. In contrast, the price promotion variable was statistically significant and had the expected sign for nearly all customers. This result was not unexpected for at least two reasons. First, few catalog consumers can observe, track, or learn average regular prices in a category over time. However, in the catalogs we study, discounts are highlighted in bold and are explicitly compared with the regular prices. Second, the regular price of an individual item never changes and the only source of variation in the average regular price is seasonal changes

Table 2   Prices

	Women's tops	Men's tops	Women's footwear
Price ($)			
Average price	25	29	51
Minimum price	18	22	33
Maximum price	31	34	67
Price promotion (%)			
1 − Average % discount	92.20	93.20	92.40
1 − Minimum % discount	99.20	99.50	99.40
1 − Maximum % discount	73.50	75.40	76.10

Notes. The price and price promotion are calculated on a monthly basis, reflecting the use of monthly transactions as the unit of observation. The findings reported in this table represent summary statistics for these monthly variables.

in the items offered. The model includes monthly dummy variables that account for seasonality and there is limited variation in the regular price variable that is independent of these seasonal effects. For these reasons, in this application, we expect customers to be more sensitive to the size of the discounts than to variations in the regular price. Because the regular price did not substantially affect customer behavior, we omitted this variable from our final model specification. For ease of exposition, we code the price promotion variable as one minus the average percentage discount. Summary statistics are shown in Table 2.

We assume that the price promotion variable captures customers' price sensitivities, which allows us to translate our model estimates into dollar values. We believe that this interpretation is reasonable as the implied price elasticities are in a range that is consistent with past research (Tellis 1988). However, we caution that the price promotion variable may capture effects other than price.

Whereas the formulation accommodates any marketing variable, we limit our analysis in this application to price. The mean utility for category j is denoted by μ_{ijt} and is specified as

$$\mu_{ijt} = \beta_{ij,0} + \beta_{i,p} p_{jt} + \gamma'_j S_t. \tag{38}$$

The p_{jt} term is the *price promotion* for category j at time t. Because this is defined as one minus the average discount, the price promotion must always be between zero and one. The intercept varies by customer and category, which allows for considerable heterogeneity in purchase rates across customers and categories. We initially estimated the model and allowed the price coefficients to vary across customers and categories (i.e., $\beta_{ij,p}$). However, the correlation of price coefficients across categories was extremely high (i.e., >0.9), suggesting that customer price coefficients were stable across categories. This evidence led us to impose the restriction that customer price sensitivity $\beta_{i,p}$ did not vary across categories. The restriction

to $\beta_{i,p}$ had little effect on the population-level price sensitivity but had the advantage of fewer positive price parameters. We caution that such a restriction may not be appropriate in all applications. Inspection of customer purchase behavior reveals evidence of seasonality: Purchases are much more likely in November and December, presumably due to the holiday season. To control for seasonality, we include a vector of month dummy variables S_t. For efficiency reasons, we assume that seasonality effects vary across categories but are homogeneous across customers.

In the regression model ΠZ, we include three customer characteristics from the U.S. Census. We have 28 zipcode-level census variables in our data set, but these variables are highly correlated. We include three variables—income, household size, and age—in the final model specification. We chose these variables because we might expect purchase and return behavior to vary with these demographics. Although the model can accommodate category characteristics, we do not include them because we only have three categories. In a broader study with many categories, including category characteristics would be appropriate.

4.   Results

In Table 3, we report the estimates for the model parameters. For expositional convenience, we omit the monthly dummies and customer intercepts from Table 3. Recall that the coefficients in Equation (37) are in a regression structure with variables that include a constant, income, household size, and age.

Table 3   Parameter Estimates

	Constant	Income	Household size	Age
Women's top intercept	−0.927 (0.017)	0.008 (0.027)	−0.570 (2.007)	−0.140 (0.246)
Men's top intercept	−0.640 (0.016)	0.007 (0.022)	−0.016 (1.812)	−0.067 (0.214)
Women's footwear intercept	−1.226 (0.020)	−0.044 (0.028)	−1.595 (1.685)	0.033 (0.259)
Price promotion	−1.597 (0.093)	0.072 (0.173)	0.056 (3.122)	1.454 (1.442)
Return cost $\bar{R}$	0.500 (0.025)	−0.102 (0.057)	1.540 (2.675)	0.271 (0.529)

	Women's tops	Men's tops	Women's footwear
Product fit σ_ψ	1.00	0.77 (0.01)	1.20 (0.016)

Notes. This table reports the posterior means for each parameter with posterior standard deviations in parentheses. Monthly dummies are included in the model but omitted from the table. The return cost ($\bar{R}$) is not category specific.

Looking down the columns of Table 3 for household size and age, we observe that neither variable is significant. We do observe weak, negative effects of income for the return cost $\bar{R}$ and the women's footwear intercept. This suggests that higher income households have lower return costs and less demand for women's shoes. The negative coefficient for the price index is consistent with a downward sloping demand function: customers purchase more frequently in the category when more items are discounted in a category. As shown in the next section of the paper, the implied price elasticity is in the range of -1.6 to -2.3, which is consistent with typical price elasticities (Tellis 1988).

The primary parameters of interest are consumer return costs R_i and the degree of product uncertainty as measured by σ_ψ. The parameter $\bar{R}$ is of the mean of the distribution of R_i for customers in our sample, and in Figure 1 we plot the distribution of estimated return costs. As shown in Figure 1, we observe considerable variation in customer return costs.

Recall that our identification strategy normalizes σ_ψ for women's tops to 1.0. We find that, relative to women's tops, customers have less uncertainty about product fit for men's tops ($\sigma_\psi = 0.78 < 1$). This is consistent with what we might expect. Women's clothing at this retailer tends to exhibit greater variation in styles and fabrics than men's clothing. There is also considerable variation in the styles and fabrics of women's clothing from year to year, whereas the changes in men's clothing are relatively small. Customers are relatively more uncertain about the fit of women's footwear than women's tops. This is also what we would expect. Because fit is much more precise for footwear than for clothing, it is generally more difficult to evaluate fit prior to physically trying the product on.
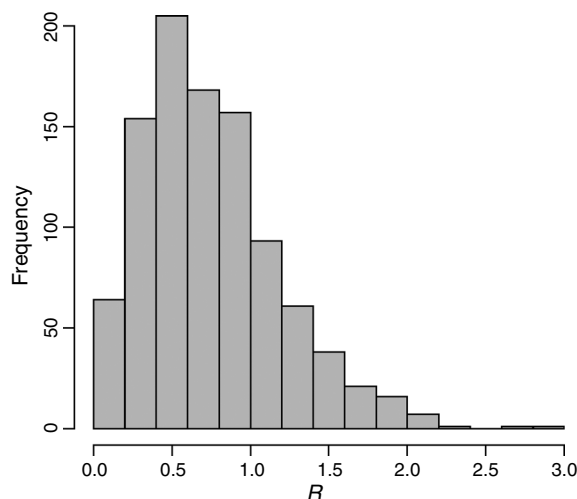
Figure 1   Histograms of Customer Return Costs (R)

Table 4   The Impact and Value of Returns

|	Women's tops	Men's tops	Women's footwear
Change in probability of ordering and keeping item each period (Δ)	0.021	0.010	0.030
Mean option value $\bar{\pi}$ ($)	5.00	3.19	15.81

Notes. The findings represent the change in the probability of a monthly order attributable to the return option (Equation (34)) and the ex ante amount that customers would be willing to pay to preserve the return option (Equation (33)).

The scale of the return cost and uncertainty parameters makes interpretation difficult. As an alternative, we report the average value of the return option both in terms of the impact on the probability that customers will order and on the change in customers' willingness to pay (see Table 4). Recall that in Equation (34), we defined Δ as the increase in the probability that a customer will place an order and keep the item when returns are allowed compared to when they are not allowed. Intuitively, this represents the increase in the probability of receiving revenue from a customer in a month that can be attributed to the return option. These probabilities can be compared directly with the base ordering probabilities in Table 1. In Figures 2(a)–2(c), we also present histograms of these estimated option values where each of the 987 customers represents a single observation for each category.

Recall that Δ represents the change in the probability of ordering and keeping an item due to the return option. Theoretically, Δ can be either positive or negative, but in this application we only observe positive values. The observed purchase probabilities in the data represent a benchmark that includes the return option. Thus, we interpret Δ as a measure of the decrease in net demand that would occur if returns were banned. From Table 1, the probabilities of ordering and keeping are 0.150 for women's tops, 0.126 for men's tops, and 0.087 for women's footwear, and the average change in this probability attributable to the return option is 0.021, 0.010, and 0.030, respectively. Given this, the opportunity to return items yields a 16% increase in demand for women's tops, a 9% increase for men's tops, and a 53% increase for women's footwear.[7] We conclude that the return option has a substantial impact on customer demand.

By scaling the customer option values $\bar{\pi}$ in a dollar metric, we can compare them with the average prices for the items reported in Table 2. As we would

---

[7] The 16% is calculated as $0.021/(0.150 - 0.021)$.

Figure 2 Histograms of Δ for Each Category

(a) Women's tops

(b) Men's tops
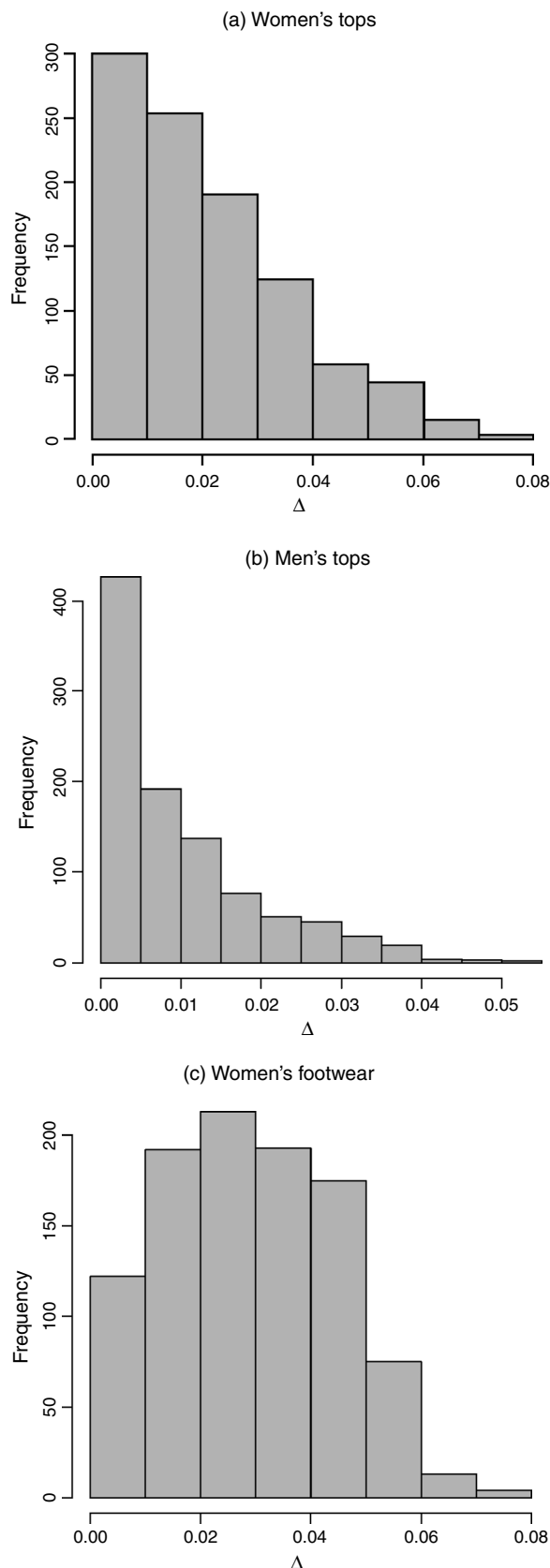
(c) Women's footwear

expect, these values are highest in the category for which there is greatest uncertainty about product fit (women's footwear). Customers would be willing to pay an average of almost $16 per transaction to have a return option in footwear but only $5 for women's tops and $3 for men's tops. These option values represent 31% of the average price of footwear, 20% of the average price of women's tops, and 11% of the average price of men's tops. Note that part of a consumer's option value of returns is implicitly consumed by hassle costs R_i. Reducing these hassle costs will substantially increase each customer's option value of returns.

4.1. Model Fit

To assess the fit of our model, we computed fit statistics both within and out of sample. To evaluate the in-sample fit, we computed simple correlations between average predicted and observed order and return rates. These correlations (not shown) were greater than 0.90 for all three categories. We illustrate this graphically in Figures A.1(b) and A.2(b) in the appendix, which depict these correlations for the women's tops category.

We also assessed our proposed model's out-of-sample predictive ability. This analysis was done "doubly" out of sample: The sample included 100 customers who were not included in the estimation of the main model parameters. In addition, for these 100 customers, only the first half of the time-series data is used to calibrate their customer-level parameters, which then are used to predict behavior in the second half of the observation period. For each category, we used two approaches to compare actual and predicted behavior for both orders and returns. First, we simply compared aggregate predictions of orders to actual orders for the 100 holdout customers in the second half of the observation period. Second, we compared actual and predicted order rates for each of the 100 holdout customers. The actual order rate is the observed order rate in the second half of the observation period, whereas the predicted order rate is the average order rate for the second half of the observation period (as predicted by the model using only data for the first half). Figures 3(a)–3(d) show these two types of plots for orders and returns in the women's tops category. Plots for the other categories are available from the authors.

The longitudinal out-of-sample aggregate predictions for orders are reasonable, although the model fails to predict some of the more extreme variations in orders. The same is true for returns, where the model accurately predicts the average level of returns over the second time period but fails to capture some of the temporary variation in returns. At the customer level,

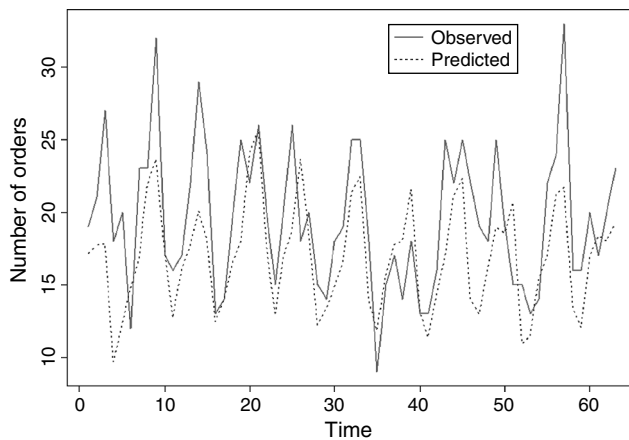**Figure 3(a)   Monthly Actual vs. Predicted Orders**
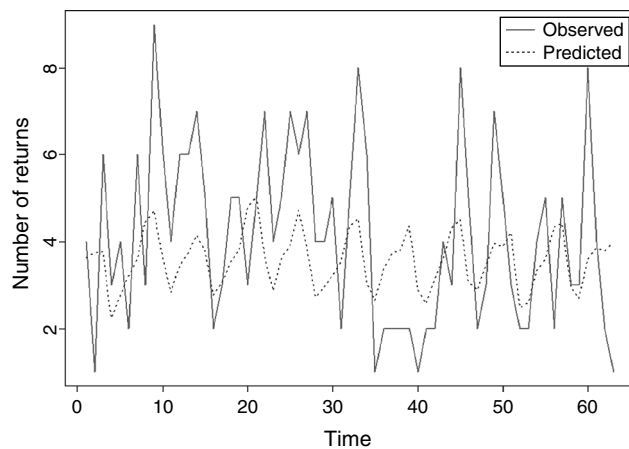

Figure 3(c)   Actual vs. Predicted Customer Orders

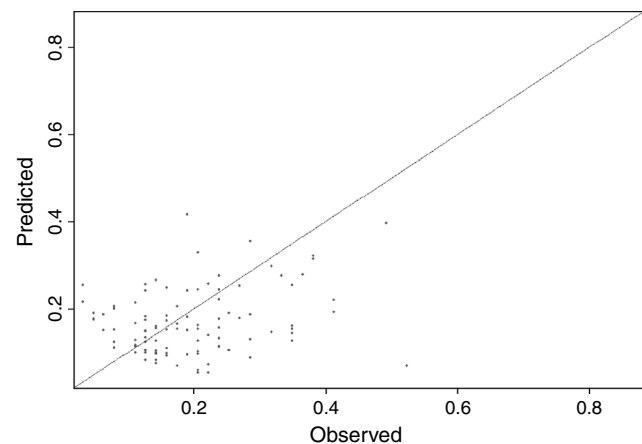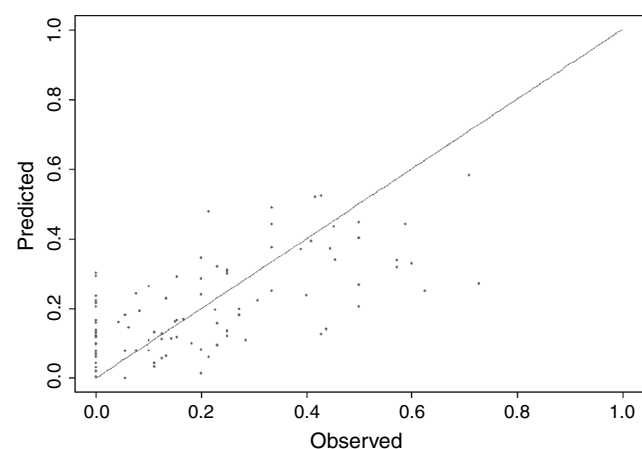**Figure 3(b)   Monthly Actual vs. Predicted Returns**

Figure 3(d)   Actual vs. Predicted Customer Returns

the model does well in predicting average order and return behavior in the second half of the observation period. Overall, the model does well in predicting the average behavior in the holdout time period—even within customer.[8] The model also does a good job in predicting temporal variation in aggregate orders but fails to capture some of the temporal variation in returns.[9]

[8] There is some evidence for the model underpredicting the number of "never returners," i.e., customers who never return products. One possible specification that can capture this is a model with a nonzero mass point at the no-return model. With our current specification, this case can only be achieved as a limit (when return costs approach infinity).

[9] We also compared our model against two benchmark models: gross demand and net demand. In the gross demand model, we considered whether a customer ordered an item; in the net demand model, we considered whether a customer ordered and kept an item. The independent variables in these benchmark models were identical to the structural model and include a customer-specific intercept, price promotion variable, and monthly dummy variables. Because the models are fitting different dimensions of the data, comparing likelihoods is not meaningful. We compared within and out-of-sample fit statistics of the benchmark models to the

5.   Implications

The previous section illustrated an application of our model to the purchase and return of apparel items. In this section, we illustrate two important implications of the model. First, we show that modeling customer return behavior can change demand estimates. Second, we show how the model may be used to optimize customer return policies.

5.1.   Demand Estimates

A key intuition from our model is that the opportunity to return an item provides customers with a valuable option. In turn, this increases each customer's demand for the product through a shift in the demand curve. We can illustrate this shift by comparing the aggregate demand curves when returns are allowed

structural model. Within sample, the fit was largely identical for the three models. As an example, Figures A.1 and A.2 compares within sample fit for the women's tops category for the three models. The out-of-sample fit of the structural model was superior to the benchmark models using a root mean squared error prediction criterion.

Figure 4   Demand for Women's Tops

(a) Net demand
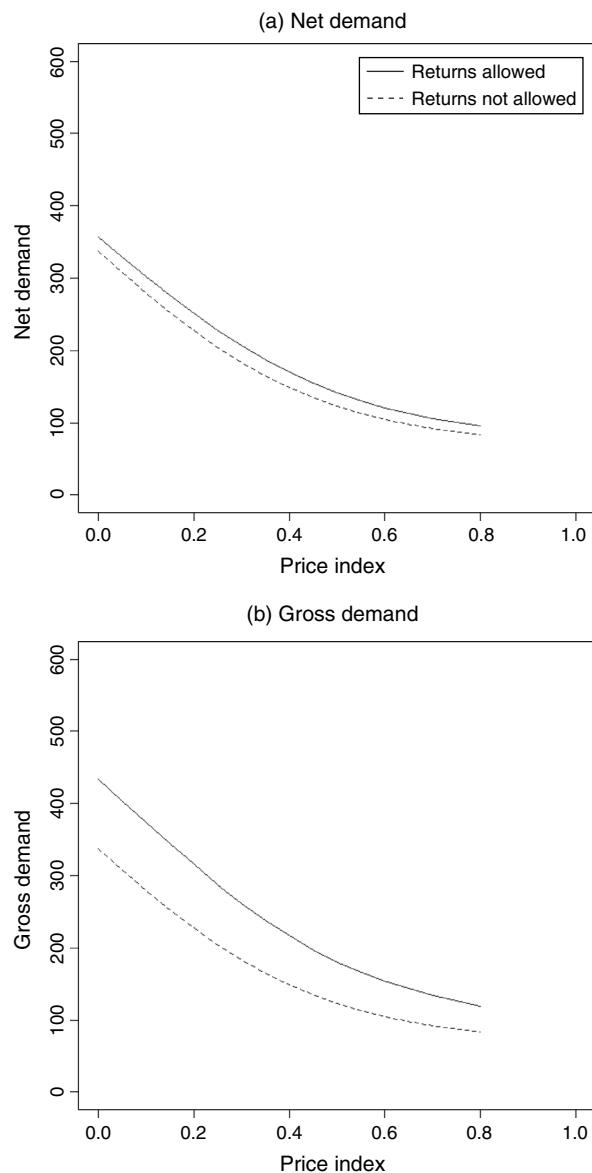
(b) Gross demand

with the same function when returns are not allowed. We compute two types of demand: gross demand and net demand. Gross demand focuses on customers' orders (ignoring returns) and net demand focuses on orders that are kept and not returned. In Figures 4(a) and 4(b), we plot these demand curves for women's tops (figures for the other categories are available from the authors).

The dashed line represents demand when returns are not allowed, and the solid line represents actual demand when returns are allowed. In our application, the opportunity to return leads to a shift in both gross demand and net demand. The shift in gross demand is considerably greater due to the large number of customers who do not purchase if they cannot return a product with poor fit. Many of these marginal customers do return their purchased items,

and so the increase in net demand is smaller than the increase in gross demand. As discussed previously, in theory the change in net demand may not always be positive. These figures also illustrate that the shift in demand is greater for low prices than for high prices. For example, in Figure 4(b), approximately 100 additional customers purchase at the lowest price but approximately 60 additional customers purchase at the highest price.

To illustrate the importance of explicitly accounting for returns when estimating demand, we compare price elasticities from our model with two benchmark demand models. In the absence of a model of customer returns, one option a researcher may pursue is to simply ignore returns and focus on gross demand. We will use this as our first benchmark model and compare it to the probability of ordering from our structural model. As a second benchmark, we consider demand net of returns (orders minus returns) and compare this with the probability of ordering and keeping in our structural model. Note that if $R_i = \infty$, the elasticity estimates from all three models are identical. But if R_i is finite, the elasticity estimates will differ because the structural model accounts for the return option. The price elasticities from each of these comparisons are provided in Table 5.

Although all three models fit customer orders similarly, this does not imply that the models yield similar elasticities. The elasticity is given by $\eta = (d\Pr/dprice)(price/\Pr)$, where $\Pr$ is the probability of an order. Because we evaluate the elasticity at the same price and the models have similar fit (i.e., $price/\Pr$ is similar), differences in the elasticities are largely attributed to the slope $d\Pr/dprice$. For the structural model, the derivatives of Equations (16) and (27) provide the respective slopes for gross and net demand elasticity.

When estimating the probability of an order (gross demand), the price elasticities estimated by the two models are reasonably similar. This was not expected but may be explained in part by the similarity of the functional form of Equation (16) $\Phi(\mu + \omega)$ and the standard choice model $\Phi(\tilde{\mu})$. The disparities in the

Table 5   Price Elasticities

	Women's tops	Men's tops	Women's footwear
Probability of an order (gross demand)			
Benchmark model	−1.87	−1.68	−2.05
Structural model	−1.84	−1.67	−2.03
Probability of an order that is kept (net demand)			
Benchmark model	−1.85	−1.65	−2.07
Structural model	−2.03	−1.78	−2.30

Table 6   Price Elasticities for Low and High Option Value Customer

	Customers with low option values			Customers with high option values		
	Women's tops	Men's tops	Women's footwear	Women's tops	Men's tops	Women's footwear
Probability of an order						
Benchmark model	−1.64	−1.50	−1.79	−1.54	−1.90	−2.47
Structural model	−1.47	−1.36	−1.55	−1.66	−2.03	−2.50
Probability of an order that is kept						
Benchmark model	−1.41	−1.35	−1.61	−2.43	−2.30	−2.53
Structural model	−1.46	−1.35	−1.53	−2.27	−2.63	−3.35

price elasticities are substantially larger when estimating the probability that an item will be ordered and kept. In the women's footwear category, accounting for the return option increases the estimated elasticity by more than 11%. On average, we find that estimating net demand without a structural model of returns leads to more inelastic demand estimates. As we will discuss next, the direction of this effect may vary.

Differences in these elasticity estimates result from customers who value the return option. In a market where customers do not value returns, we would expect few differences between the structural and benchmark models. In contrast, if all customers value returns then we would expect to see large disparities in the estimated elasticities. To illustrate this point, we consider the 5% of customers with the lowest and 5% of customers with the highest option values in each category. Similar to Table 5, we compute elasticities for the benchmark models and our structural model. These estimates are reported in Table 6.

For customers with low option values, the gross demand model yields elasticities that are 9%–13% more elastic than the structural model. However, for customers with high option values, we observe the opposite result: the gross demand model yields elasticities that are 1%–8% less elastic. Comparison of the elasticities from the net demand model illustrates a similar pattern: price elasticities vary and these differences can be large. For example, among customers with high option values, the price elasticity from the women's footwear structural model is 30% more elastic than estimates from the net demand model.

We conclude that there may be substantial differences between the elasticities of the structural model and the benchmark models, and these differences may be positive or negative. If customers consider the option of returning in their purchase decisions, than it is appropriate to account for this in the demand model. Because our structural demand model explicitly accounts for this possibility, it may be argued that this model will yield more accurate price elasticities.

5.2. Optimizing Return Policies
The structural nature of the model allows us to make predictions about how customers will respond if a firm varies its return policy. By varying a return policy, a firm can raise or lower customers' return costs. For example, rather than charging customers for the cost of the return postage, the company could lower customers' returns costs by paying for some or all of the postage costs (as Macy's does for its elite and platinum cardholders). Alternatively, the firm could raise the return costs by charging an administrative (restocking) fee for accepting returns.

We investigate the impact of return policies by computing counterfactuals using different return costs. In particular, we consider return policies of the form:

$$R_i^{\text{new}} = K\widehat{R}_i, \quad i = 1, \dots, N. \quad (39)$$

The benchmark return policy is $K = 1$, under which all customers face the return costs directly estimated from the data. As we increase or decrease K, we can estimate the resulting reduction in the number of orders received and the change in the number of returns. This in turn allows us to estimate the impact on firm profits.[10]

To calculate firm profits, we first define the gross profit earned from category j with return policy α as

$$\Pi_j(K) = m_j \sum_{i=1}^{N} \Pr(U_i^O > 0, U_i^K > 0 \mid R_i^{\text{new}}, \widehat{\beta}_{ij})$$

$$- c_j \sum_{i=1}^{N} \Pr(U_i^O > 0, U_i^K < 0 \mid R_i^{\text{new}}, \widehat{\beta}_{ij}). \quad (40)$$

In this expression, we use m_j to denote the average profit margin of items in category j and c_j as the firm's cost if an item in category j is returned. This cost includes both the administrative cost of processing the return transactions and the depreciation in the value of the returned item. In the absence of any information about the actual value of c_j, we assumed for the purpose of this illustration that it was equal to 35% of m_j. For simplicity, we also assumed that varying customers' return costs only affects the firm's

---

[10] Note that we have kept K constant across customers. This implies that the policies we consider are return policies, where all customers are hit by the same percent increase or decrease in return costs. It is straightforward to calculate the profits from segmented or even individual return polices by allowing K to vary across customers.

expected profits through changes in the volume of orders and returns. In practice, some of the options available to change the cost of a return, such as waiving the cost of return postage or charging a restocking fee, could also directly impact the firm's profit function through c_j. Modifying (40) to accommodate these changes is straightforward.

In Figures 5(a)–5(c), we illustrate the change in the expected profits (Π) as a function of K for the three product categories. For the women's tops category, the current return policy (i.e., $K = 1$) appears to be close to optimal, although this finding depends upon our somewhat arbitrary assumption about c_j. Of greater interest is the fact that for men's tops, it is optimal to make returns more expensive compared to women's tops. Taken literally, our estimates suggest that the optimal policy for men's tops is to not allow returns (i.e., $K = \infty$). A more conservative interpretation of the results is that the return policy for women's tops should be more lenient than that for men's tops. For women's footwear, our model suggests that it is optimal to make returns less expensive (i.e., $K < 1$).

These comparisons make intuitive sense. Recall from Table 2 that men's tops had the lowest uncertainty and customers placed less value on the option of returning items in this category (compared to the other two categories). In contrast, women's footwear had the highest uncertainty and customers place greater value on their ability to return these items. Tightening the return policy will have less impact on demand for men's tops than on demand for footwear, suggesting that it is optimal to offer different return policies across these categories.

To illustrate how managers can translate K from a utility metric to a dollar metric, we compare the observed value $\pi(R)$ with the optimal value $\pi(KR)$. Recall that for women's footwear, customers were willing to pay almost \$16 for the option of returning. The optimization results suggest that return costs should be lowered to a level that increases the option value of returns by 7%. This is equivalent to providing customers with \$1 in option value, which might be achieved by subsidizing the cost of return postage, for example. The result also suggests that offering free return postage for returned items may be too generous and may reduce profits.

We caution that these counterfactual simulations do not account for competitive reactions and this may affect the optimal return policy. Whether we overstate or understate the change in profit depends on the nature of the competitive reaction. For example, assume the focal retailer increases the return costs for women's tops and competitors respond by increasing their return costs. This may soften overall competition and our simulation may underestimate the change in

Figure 5   Optimal Return Policies for Each Category

(a) Women's tops

(b) Men's tops

(c) Women's footwear
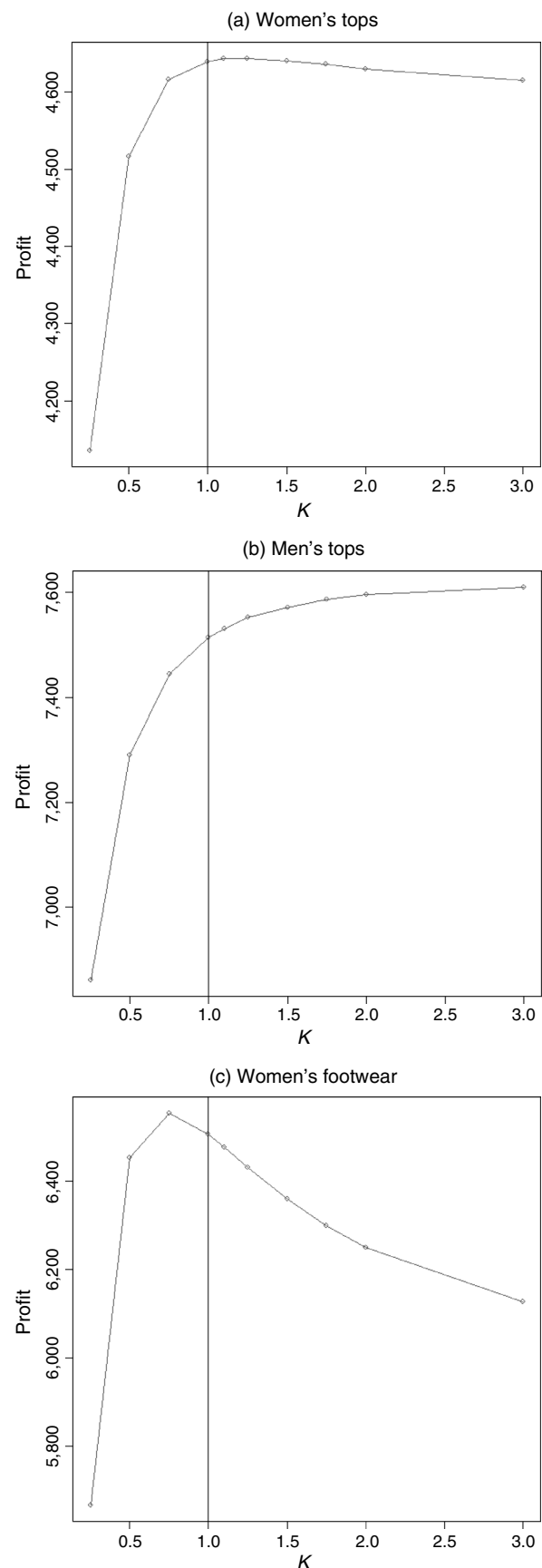
profit. In contrast, the opposite result may occur if return costs are decreased.

Although it is important to assess competitive reactions in return policies, this cannot easily be done through empirical analysis. In our data, there are no substantive changes in the retailer's return policy for a 10-year period and there is no variation in return policies among categories. Because return policies are changed infrequently for a single retailer, a researcher could obtain data from multiple retailers and rely on variation in return policies across retailers. In the Technical Appendix (available at http://mktsci.pubs.informs.org), we show how the full model can be extended to incorporate competition. We also illustrate how to incorporate changes in competitors' return policies using the outside option in a partial equilibrium model. However, neither approach is empirically feasible in our application. In the industry we analyze, we are not aware of a multiretailer database nor is there variation in competitors' return policies. Given these limitations, it may be more useful to assess the impact of competitive response with qualitative, managerial insights.

If the simulations are correct, we would expect to see them reflected in the actual return policies that firms use in practice. For example, we might expect fewer restrictions on returns of footwear than on returns of apparel. As a preliminary test of this prediction, we identified a sample of remote retailers who specialize in either women's clothing, men's clothing, or footwear. We used a list provided by Google.com that includes the names and Web addresses for a large set of remote retailers in each of these categories. We then visited each retailer's website and recorded their return policies.

Our final sample included 44 footwear retailers, 22 men's clothing retailers, and 46 women's clothing retailers. The low number of men's clothing retailers reflects the relative paucity of remote retailers specializing in men's clothing. Within this sample of retailers, we measured the time limits imposed on the ability to return items (measured in days after delivery date). The averages are very similar for men's and women's apparel: 22 days for women's clothing and 24 days for men's clothing. The median number of days allowed is 14 for women's clothing and 15 for men's clothing. In contrast, footwear retailers have much more liberal return policies. The mean number of days allowed for customer returns is 45, and the median number of days is 30. The difference in the number of days allowed for customer returns is significantly different ($p < 0.01$) for footwear compared to men's and women's clothing. We conclude that the evidence is at least partially consistent with the predictions provided by the model. Although we did not observe a difference in return policies for men's and women's apparel, there is evidence that footwear retailers tend to have more liberal return policies than apparel retailers.

6. Conclusions

Numerous retailers offer customers the option to return previously purchased merchandise. Although this provides customers with an option that has measurable value, this value has not been measured. In this paper, we develop a structural model that generalizes a standard purchase incidence model and incorporates a consumer's decision to purchase and return an item. We then empirically quantify the option value each consumer enjoys due to a firm's return policy. We illustrate how the model can be used by a retailer to optimize its return policies for different product categories.

To estimate the model, we had to overcome problems with both tractability and identification. The tractability issue was overcome by replacing the purchase probability expression with a more flexible function that allowed a direct comparison between this more general model and the standard purchase incidence model. The identification problem arises because customers' individual return costs are confounded with their uncertainties about product fit. We disentangle these two factors by recognizing that a customer's return costs are often fixed across different product categories, while the uncertainty about fit may vary. By estimating the model across multiple categories, we are able to separately identify the return costs and the category-specific uncertainty parameters.

We illustrate the information that can be learned from the model by applying it to a sample of data provided by a mail-order apparel company. We use a sample of 987 customers who had all made purchases of women's tops, men's tops, and women's footwear from the firm. We find that there is substantial variation in option values across customers and that the impact of the return option on demand is substantial for many customers. In addition, the findings reveal that customers are generally more uncertain about product fit for women's clothing than for men's clothing and that there is even more uncertainty about fit for women's footwear. As a result, customers place the highest value on the option of returning women's footwear and the lowest value on returning men's clothing. Our results show that the opportunity to return items leads to an average increase in demand of 16% for women's tops, 9% for men's tops, and 53% for women's footwear.

We also illustrate how varying the cost of returning an item affects a firm's profits. As the earlier results suggest, we find that optimal return policies are more lenient for women's footwear than for clothing. We would expect to see this pattern reflected in

the actual return policies that firms use in practice. To evaluate this prediction, we surveyed the return policies at a sample of specialty footwear and apparel retailers. The findings reveal that return policies are generally more lenient for footwear than for apparel, although we observed no differences in return policies for men's and women's apparel.

This is the first attempt to model customer return decisions in a structural framework. There are many opportunities to extend our model in future research. This includes investigations of customer learning about product fit, exploring cross-channel differences in return behavior, allowing for concavity in return costs, allowing for temporal or cross-category dependence in demand and returns, and incorporating retail competition.

Acknowledgments

Appendix

Figure A.1   Predicted vs. Observed Net Demand for Women's Tops

(a) Net orders: Benchmark model

(b) Net orders: Structural model
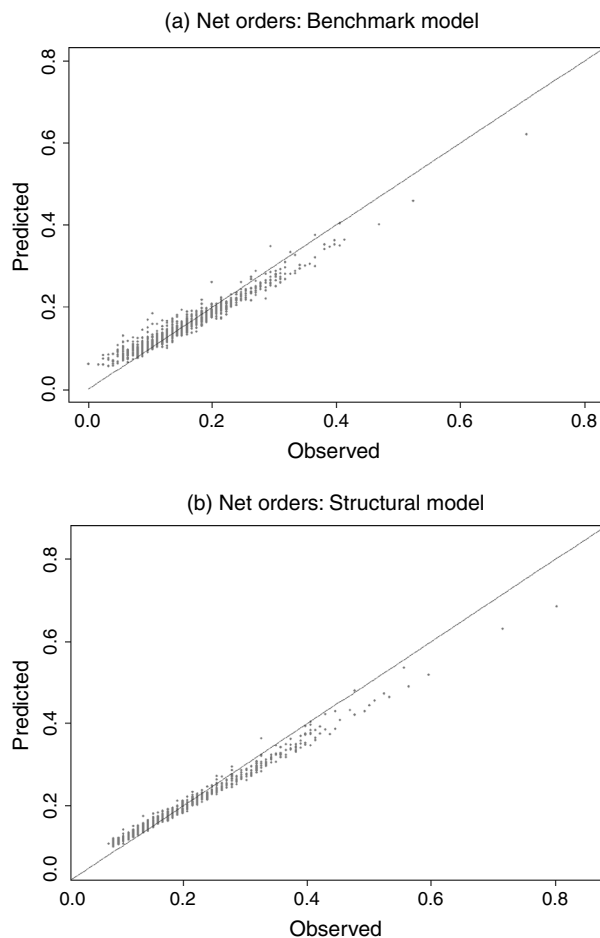
Figure A.2   Predicted vs. Observed Gross Demand for Women's Tops

(a) Gross orders: Benchmark model

(b) Gross orders: Structural model
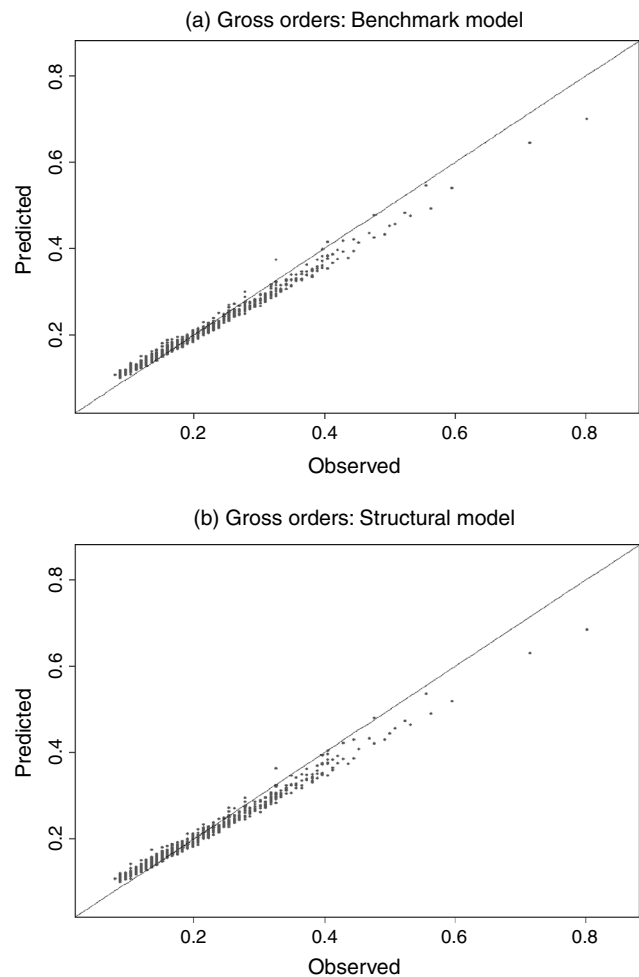
References

Alba, J. W., S. Broniarczyk, T. Shimp, J. E. Urbany. 1994. The influence of prior beliefs, frequency cues, and magnitude cues on consumers' perceptions of comparative price data. *J. Consumer Res.* **21**(2) 219–235.

Anderson, E. T., G. J. Fitzsimons, D. Simester. 2006. Measuring and mitigating the costs of stockouts. *Management Sci.* **52**(11) 1751–1763.

Anderson, E. T., K. Hansen, D. I. Simester, L. K. Wang. 2009. How price affects returns: The perceived value and incremental customer effects. Working paper, Kellogg School of Management, Northwestern University, Evanston, IL.

Blanchard, D. 2005. Moving forward in reverse. *Logist. Today* **46**(7) 1.

Blattberg, R. C., R. Briesch, E. J. Fox. 1995. How promotions work. *Marketing Sci.* **14**(3) G122–G132.

Che, Y.-K. 1996. Customer return policies for experience goods. *J. Indust. Econom.* **44**(1) 17–24.

Davis, S., E. Gerstner, M. Hagerty. 1995. Money back guarantees in retailing: Matching products to consumer tastes. *J. Retailing* **71**(1) 7–22.

Davis, S., M. Hagerty, E. Gerstner. 1998. Return policies and the optimal level of hassle. *J. Econom. Bus.* **50** 445–460.

Emmons, H., S. M. Gilbert. 1998. The role of returns policies in pricing and inventory decisions for catalogue goods. *Management Sci.* **44**(2) 276–283.

Heiman, A., B. McWilliams, D. Zilberman. 2001. Demonstrations and money-back guarantees: Market mechanisms to reduce uncertainty. *J. Bus. Res.* **54** 71–84.

Hess, J. D., G. E. Mayhew. 1997. Modeling merchandise returns in direct marketing. *J. Direct Marketing* **11**(2) 20–35.

Ho, T. H., C. S. Tang. 2004. Introduction to the special issue on marketing and operations management interfaces and coordination. *Management Sci.* **50**(4) 429–430.

Jedidi, K., C. F. Mela, S. Gupta. 1999. Managing advertising and promotion for long-run profitability. *Marketing Sci.* **18**(1) 1–22.

Kahneman, D., A. Tversky. 1984. Choices, values and frames. *Amer. Psychologist* **39** 341–350.

Karmarkar, U. S. 1996. Integrative research in marketing and operations management. *J. Marketing Res.* **33**(2) 125–133.

Kiesmüller, G. P., E. A. van der Laan. 2001. An inventory model with dependent product demands and returns. *Internat. J. Production Econom.* **72**(1) 73–87.

Kulp, S. C., H. L. Lee, E. Ofek. 2004. Manufacturer benefits from information integration with retail customers. *Management Sci.* **50**(4) 431–444.

Lichtenstein, D. R., W. O. Bearden. 1989. Contextual influence on perceptions of merchant-supplied reference prices. *J. Consumer Res.* **16** 55–66.

Merrick, A., I. Brat. 2005. Taking back that blender gets harder—Sears is the latest retailer to tighten returns policy; how to avoid being refused. *Wall Street Journal* (December 15) D1.

Mohl, B. 2003. For duo, basement ban is no bargain. *Boston Globe* (July 13) C1.

Moorthy, S., K. Srinivasan. 1995. Signaling quality with a money-back guarantee: The role of transaction costs. *Marketing Sci.* **14**(4) 442–466.

Padmanabhan, V., I. P. L. Png. 1997. Manufacturer's returns policies and retail competition. *Marketing Sci.* **16**(1) 81–94.

Pasternack, B. A. 1985. Optimal pricing and return policies for perishable commodities. *Marketing Sci.* **4**(2) 166–176.

Rogers, D. S., R. S. Tibben-Lembke. 2001. An examination of reverse logistics practices. *J. Bus. Logist.* **22**(2) 129–248.

Savaskan, R. C., S. Bhattacharya, L. N. Van Wassenhove. 2004. Closed-loop supply chain models with product remanufacturing. *Management Sci.* **50**(2) 239–252.

Shulman, J., A. Coughlan, C. Savaskan. 2009. Optimal restocking fees and information provision in an integrated demand-supply model of product returns. *Manufacturing Service Oper. Management.* Forthcoming.

Tellis, G. J. 1988. The price sensitivity of competitive demand: A meta-analysis of sales response models. *J. Marketing Res.* **15**(3) 331–341.

Walker, A. 2004. Retailers are getting tough on merchandise returns; trying to reduce costly fraud. *Baltimore Sun* (December 26) A10.