Why Outlet Stores Exist:
Averting Cannibalization in Product Line Extensions

Donald Ngwe*

July 31, 2014

Abstract

Many firms sell goods through outlet stores in addition to regular stores, particularly in the fashion industry. Outlet stores offer attractive prices at locations far from central shopping districts. The main perspectives as to why outlet stores exist can be broadly classified into inventory management, geographic segmentation, and price discrimination through consumer self-selection. I evaluate these perspectives in the context of a major fashion goods firm using newly available and highly granular data. Model-free evidence suggests that inventory management and geographic segmentation are not the main drivers for outlet store use. Consumers who shop at outlet stores also do not differ significantly from those who shop at regular stores in terms of income. I use a structural demand model to show that consumers are considerably heterogeneous in their sensitivity to travel distance and taste for product newness. I then develop a supply model to predict product development responses to changes in store locations. Through policy simulations, I find that the firm uses outlet stores to serve lower-value consumers who self-select by traveling to outlet stores from central shopping districts. The firm sells older, less desirable merchandise through outlet stores to prevent cannibalization of regular store revenues by means of exploiting the positive correlation between consumers' travel sensitivity and taste for new products. I find that the rate of product introduction in regular stores would fall by 13% if outlet stores were closed down, while variable profits would decline by 19%.

1 Introduction

Outlet stores are a fixture of the American retail landscape. These are brick-and-mortar stores that offer deep discounts in locations far away from most consumers. Firms operate outlet stores in addition to regular stores, which are located in central shopping districts.

*Email: dngwe@hbs.edu. This paper is based on my dissertation. I am grateful to my committee members Chris Conlon, Brett Gordon, Kate Ho, Mike Riordan, and Scott Shriver for their guidance and encouragement. I also thank Yeon-Koo Che, Jonathan Dingel, Jessie Handbury, Oded Netzer, Bernard Salanié, Patrick Sun, Olivier Toubia, and participants of Columbia's industrial organization colloquium for helpful comments and suggestions. All errors are my own.

Outlet stores operated by different firms are often agglomerated in sprawling outlet malls off interstate highways. In 2012 there were 185 outlet malls in the US, which generated an estimated $25.4 billion in revenues (Humphers, 2012).

There are several perspectives on why outlet stores have become a widely adopted selling strategy. The first is *inventory management*: outlet stores provide firms with a cost-efficient way to dispose of excess inventory. The second is *geographic segmentation*: outlet stores cater to lower-value consumers that reside around outlet malls. The third is *consumer self-selection*: lower-value consumers travel greater distances to avail of discounted products.

In this paper, I evaluate the relevance of each of these proposed explanations to the case of a major fashion goods firm with a heavy outlet store presence. Using new and highly granular data, I am able to observe both inventory flows between store formats, and locations and sales records of individual consumers—rich sources of model-free evidence. I then make use of structural models of demand and supply to predict consumer behavior and firm product decisions under counterfactual store configurations.

It is evident from observing product flows alone that inventory management is not an essential function of the firm's outlet stores. The firm sells a significant fraction of units of each style through the outlet channel. It is also immediately clear that outlet stores do not primarily serve the communities in their vicinity—most of each outlet store's revenues are attributed to consumers for whom a regular store is closer to home. This suggests that the firm's main motivation for operating outlet stores might be to price discriminate among its consumers by forcing the most price-sensitive among them to travel to obtain discounts.

Surprisingly, consumers who shop at outlet stores do not differ significantly from consumers who shop at regular stores in terms of observable characteristics such as income. They make purchases at roughly the same frequency, and have had about the same time elapse since their first purchase of the brand. Taking these factors into account, I propose a demand model that characterizes how consumers make their purchase decisions. I use the demand model to estimate the extent to which consumers vary in their unobservable characteristics, and to show that outlet store consumers differ from regular store consumers in two ways: their sensitivity to travel distance and their taste for product newness. In addition, I find a strong positive correlation between these two values.

I hypothesize that the firm exploits the positive correlation between consumer travel sensitivity and taste for new products by selling older products in its outlet stores. I test this notion by setting the correlation to zero and simulating the corresponding purchase behavior. I find that the resulting advantage to operating outlet stores is much diminished, owing to the fact that outlet stores would cannibalize a larger portion of regular store revenues.

In order to better characterize the consumer's choice set in the absence of outlet stores, I build a supply model in which the firm optimally sets prices and product introduction rates given store locations. While prices can be adequately modeled using a standard monopoly pricing assumption, modeling the firm's product choice presents a nontrivial challenge. I address the problem by developing a probabilistic model of product choice. Rather than requiring the firm to choose characteristics individually for each of hundreds of products, I describe the firm's choice set in terms of a joint probability distribution of characteristics.

The firm's problem can then be reduced to choosing the parameters of this distribution. Since product ages are of particular importance to consumers, I focus on the firm's choice of the rate of product introductions and reassignment to outlets, which are arguably the components of product quality over which the firm has the highest degree of control.

I find that the firm is able to serve a much narrower range of consumers in the absence of outlet stores. With only its regular distribution channel available, the firm would expand its regular retail audience by lowering prices and the rate of product introduction (and hence the average age) of its products, but would be unable to attain the same level of coverage without the geographic differentiation enabled by outlets. This reveals an additional benefit of having outlet stores: they enable the firm to increase its rate of product introduction in the regular format. I find that the firm introduces 13% more new styles with dual distribution than with only regular stores.

The paper proceeds as follows. I review the related literature in Section 2. In Section 3, I describe the data. In Section 4, I provide preliminary evidence of how outlet stores work. In section 5, I outline the demand model I use to estimate preferences, discuss the estimation procedure, and present the estimates. In Section 6, I outline the supply model I use to describe firm product choice, and present the implied marginal and product development costs. In Section 7, I perform policy simulations that highlight the benefit of operating outlet stores. Section 8 concludes.

2 Related literature

This work contributes to several literatures in marketing and economics. It is the first empirical paper to study the incentives behind outlet store retail. It builds on existing work on product line decisions. It proposes a technique to model endogenous product choice for cases in which a large number of products comprise each product line. The underlying structure of the firm's problem that I model belongs to the class of multidimensional screening models, for which few general results are available and no empirical work has been performed. Finally, the paper demonstrates that outlet stores allow the firm to improve quality in its regular stores, which may countervail brand dilution.

Several theories exist about how why firms build and sell goods through outlet stores. Deneckere and McAfee (1996) derive conditions under which a firm would damage or 'crimp' a portion of its goods to increase profits by expanding its market share, and put forth outlet stores as an example of such a damaged goods strategy. This paper picks up this example and provides the first empirical demonstration of a successful damaged goods policy. Coughlan and Soberman (2005) show that dual distribution (i.e. having both regular and outlet stores) is more profitable than single channel distribution when the range of service sensitivity is low relative to the range of price sensitivity. While these sensitivities were independent in their model, I look chiefly at how the correlation between consumer sensitivities matters. In recent empirical work, Qian et al. (2013) show that the opening of an outlet store had substantial positive spillovers for a retailer's regular channel, and ascribe this spillover to the advertising effects of a new store opening. This paper goes further by studying how the firm's optimal product line choices are influenced by its store locations—thereby offering an alternative mechanism by which outlet stores can have positive spillovers.

More generally, this paper offers a new point of view on how product lines should be designed to effectively segment consumers. Previous work on product line design has explored the benefits of broadening product lines (Kekre and Srinivasan 1990; Bayus and Putsis 1999), methods for selecting optimal product lines (Moorthy 1984; Green and Krieger 1985; McBride and Zufryden 1988; Dobson and Kalish 1988; Netessine and Taylor 2007), cannibalization between product lines (Desai 2001), pricing (Reibstein and Gatignon 1984; Draganska and Jain 2006), and the effects on brand equity of product line extensions (Randall, Ulrich, and Reibstein 1998). My paper contributes to this body of work by demonstrating the importance of accounting for the full extent of consumer heterogeneity in making product line decisions. It also shows how concerns like cannibalization can be ameliorated by a careful design of product line attributes.

I develop an explicit model of product line choice that corresponds to the institutional details of the fashion goods industry. The large number of products in each product line poses a particular challenge. While existing work has modeled endogenous product choice for a single multidimensional good (Fan 2010) or for several single-dimensional goods (Draganska, Mazzeo, and Seim 2010; Crawford, Shcherbakov and Shum 2011), none has addressed the product choice problem of a firm with several multidimensional products. I introduce a simple and tractable method of describing this choice. Modeling the firm as choosing the parameters of a distribution of product characteristics, rather than the characteristics that make up each individual product, dramatically reduces the number of choice variables for the firm. It may also be a more realistic representation of decision-making in many sectors.

The importance of allowing product design to be endogenously determined in equilibrium has been emphasized in many recent papers. Kuksov (2004) shows that firms may respond to lower buyer search costs by increasing product differentiation and thus diminishing price competition. There are many other instances in which allowing for endogenous product differentiation changes the sign of welfare effects.

This paper's central premise is that the choice of whether to open outlet stores and what to stock them with is a type of multidimensional screening problem. Empirical models of multidimensional product choice are particularly useful because they can be used to complement lessons from theoretical work in multidimensional screening. The obstacles to obtaining general results in multidimensional screening are well-documented by Rochet and Stole (2003). Full solutions to this problem are available for the discrete two-type case (Armstrong and Rochet 1999) and other cases for which the form of consumer heterogeneity is severely restricted (e.g. Armstrong 1996). It is difficult to see how these models' predictions would manifest in actual product decisions, such as those in my empirical setting. By using demand and supply models that are not anchored to any particular screening model, I am able to provide evidence for the applicability of existing results to real world settings and the significance of multidimensional screening for firms in general.

3 Data and industry background

The first outlet stores appeared in the Eastern United States in the 1930s. These stores were attached to factories and sold overruns, irregulars, and slightly damaged goods. Outlet stores initially catered to only the firm's employees, but the stores' market audience quickly expanded to include regular consumers. Until the 1970s, firms continued to use outlet stores primarily to dispose of excess inventory, even as they established them independently of manufacturing centers.

The modern outlet store has evolved into a considerably different format from its earlier incarnations. In many ways, outlet store goods now constitute distinct product lines, rather than mere excess inventory. Many firms design products exclusively for sale in outlet stores (though they may prefer to limit awareness of the practice among consumers). Revenues from outlet stores often rival, and sometimes exceed, revenues from a firm's regular retail formats.

One feature of the outlet store that remains unchanged is its distance from central shopping districts. In fact, an entire industry of outlet mall operators owes its existence to the prevalence of this selling strategy among clothing and fashion goods retailers. The practice of selling goods in hard-to-access locations would seem curious were it not so common. Deneckere and McAfee (1996) provide a relevant argument in this regard by showing that a firm may profit from 'damaging' a portion of its goods. They also point out that the practice is widespread: certain slower microprocessors, student editions of software, and outlet store offerings can all be considered damaged goods.

Yet many firms choose not to sell through outlet stores; adoption is variable even within narrowly-defined categories. For instance, premium apparel brands Brooks Brothers, Hugo Boss, and Ralph Lauren have several outlet store locations, but Chanel, Burberry, and Zegna have few or none. Coughlan and Soberman (2005) provide an explanation for this fact that rests on the form of consumer heterogeneity. They show that firms find single-channel distribution superior when the range of service sensitivity among consumers is high relative to the range of price sensitivity.

The category to which our firm belongs generates annual revenues of about $9 billion in the US. Our firm is the market leader, with between 30 and 40 percent market share. About 60% of the firm's revenues are sourced from sales of its main category.

The data used for this study consists of transaction-level records from July 2006 to March 2011. The sample includes all purchases of products made by US consumers in firm-operated channels. Excluded from this sample are online and department store sales, which according to the firm's managers accounted for less than 10% of total revenue.

The firm is able to track repeat purchase behavior by consumers. Available information on consumers includes their billing zip codes, date of first purchase at a store, and their total lifetime expenditures on the firm's products. Each record contains detailed information on the consumer, the product, and the store. Product attributes include color, silhouette, materials, collection, release date, and a code that uniquely identifies each style. Store attributes include their location, weeklong foot traffic, and format type.

For the analysis in this paper, I focus on main category purchases in physical stores. While this excludes a considerable number of non-main category purchases, those observations are used to proxy for the number of consumers who visit a store but do not make a main category purchase.

The firm's overall distribution strategy is fairly typical among brands with outlet store locations. The firm introduces most of its new products in its regular stores, which are located in central shopping districts. After a few months, these products are pulled out of the regular stores and transferred to the outlet stores. The firm also produces styles that are sold exclusively in outlet stores.

Table 1 summarizes the differences between the firm's two store formats. The most obvious difference is in price: the typical product goes for about $300 in regular retail stores, while most outlet store products sell for less than half that price. Outlet stores are also bigger than regular retail stores in terms of square footage and the number of styles on shelf; however, each market is typically served by several regular retail stores and a single outlet store.

Table 1: Average Store Characteristics

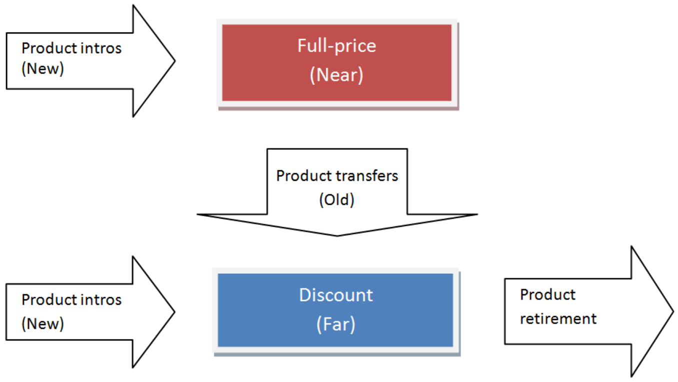
Store format:	Regular	Outlet
Transacted price	299	126
Number of products on shelf	150	432
% premium material	29.3	24.4
% basic material	35.8	39.3
Months since product intro	11.9	15.1
Square footage	2,718	4,536
Weekly foot traffic	2,845	7,677
Annual revenue	1,267,480	5,048,774
Total revenue in format	461M	722M
Total number of stores	367	143

The composition of available product choices in the two formats do not differ greatly according to stylistic characteristics. Most products are made of either a basic or a premium material and the two formats carry about the same percent of each type. Where the assortments do differ greatly is in age—time that has elapsed since the products were introduced. A fashion good's age is likely an important determinant of its attractiveness in an industry that is marked by constant product updating.

4 Preliminary evidence

In this section I use a descriptive analysis of the data to offer preliminary evidence of the value to the firm of having outlet stores. In each subsection, I provide model-free evidence

Figure 1: Product flows



that speaks to each of three main possible purposes: inventory management, geographic segmentation, and consumer self-selection.

4.1 Inventory management

I first consider the relevance of outlet stores in managing the firm's inventory: particularly the disposal of excess supply. This purpose serves as the historical basis for outlet stores' emergence, and continues to be relevant for many firms. As I show in the following discussion, however, inventory management does not appear to be a regular purpose of the firm's outlet stores.

At the most basic level, the firm manufactures two types of products, which I term *original* and *factory*. Original products are introduced in the regular stores, and after a few months, taken out of regular stores and sold in outlet stores. Factory products are sold only in outlet stores. Figure 1 summarizes these flows. At any given point, an outlet store offers about as many original products as factory products. While original products are typically thought of as more desirable than factory products, anecdotal evidence suggests that consumers are seldom able to distinguish one from the other, or even aware of the distinction. Table 2 contains information on the flows of these product types.

Table 2: Inventory Flows

Product type:	Original	Factory
Average styles introduced per year	336	132
Average months sold in regular format	10.32	N/A
Average months sold in outlet format	6.97	11.03
Average total units per style sold in regular format	3,103	N/A
Average total units per style sold in outlet format	2,707	12,725
Average composition of styles in outlet format (%)	42.71	57.29

Inspecting product flows alone suggests that the firm does not use outlet stores for the traditional purpose of disposing of excess inventory. First, it is not the firm's policy to sell defective merchandise in either of its channels. Second, the firm manufactures a product line that is meant for exclusive sale in its outlet stores. And third, close to half of the units of each style that is introduced in regular stores is sold in outlet stores. This implies that the life of 'original' products in outlet stores represents a deliberate aging strategy rather than a dump of excess inventory.

4.2 Geographic segmentation

Given how the firm uses location to distinguish each product line, a natural hypothesis is that outlet stores are designed to segment consumers according to geography. In fact, outlet stores are located in areas that have lower population density and lower income than the areas around regular stores.

Table 3 catalogues average consumer characteristics in each format that are observable in the data. The averages are taken over all purchases in each format. Median household incomes by zip code from the 2010 American Community Survey are used to proxy for a consumer's income. The consumer's travel distance is the distance between centroids of the consumer's billing zip code and the store's zip code.

Noteworthy in Table 3 is the absence of an appreciable difference in observable characteristics between consumers who buy from the two formats. They resemble each other not only in income, but also in their level of experience with the firm's products. As will be clear from the succeeding discussion, the difference in travel distance reflects the fact that consumers in both formats live in the same areas, but must travel farther to access outlet stores.

Table 3: Average Consumer Characteristics by Store Format

Store format:	Regular	Outlet
Income	71,231	65,226
	(27,780)	(23,670)
Years since first purchase	2.51	2.25
	(3.40)	(3.08)
Lifetime expenditures on brand	4,071	2,475
	(12,757)	(11,425)
Travel distance	9.53	20.44
	(8.49)	(15.65)

Standard errors are in parentheses.

An alternative way of thinking about classes of consumers is presented in Table 4. In this table, I consider consumers who have made at least two purchases in the sample and group them according to the store formats at which they made the transactions. Consumers either shopped at exclusively one format, or at both formats. *Share* refers to what percent of all consumers belongs to each class. *Outlet closer* is the percent of each class of consumers for whom the closest store is an outlet. The main takeaway from Table 4 is that even within the class of consumers who shop exclusively at outlet stores, 70.5 percent live closer to regular stores.

Table 4: Consumer behavior

	Only regular store	Only outlet store	Multi-home
Share (%)	13.2	56.8	30.0
Outlet closer (%)	7.0	29.5	15.3
Average income ($)	73,489	64,554	69,294

I take a core-based statistical area (CBSA) to be a reasonable geographic market definition.[1] I choose months as a temporal market definition. While perhaps a shorter time period than actual consumers take to return to the market, rapidly changing choice sets necessitate a tightly defined market period. Table 5 has descriptive statistics for the average market according to my definition.

Table 5: Average Market Characteristics

	Mean	St Dev
Number of regular stores	1.96	3.52
Number of outlet stores	0.66	0.66
Revenue	186,666.00	371,510.50
Market size (#consumers)	92,870.51	186,769.80

A market is a CBSA-month.

Figure 2 identifies where the firm's consumers live in Indianapolis, Indiana and shows the market population density by zip code. Indianapolis is a typical market for the firm, which it serves with two regular store locations and one outlet store. For the purposes of this figure, a 'consumer' is an individual who purchased at least one item from the firm within the five-year sample.

Figure 2 highlights the fact that the outlet store is located in an area where very few consumers reside. This agrees with what is found in the national sample, where there is a relatively small group of consumers for whom the closest store is an outlet store.

Figure 3 shows from where revenues at the outlet store are sourced. The shading of the regions closely resembles the market population density shown in Figure 2. Most of the outlet store's revenues are attributed to consumers who live in the central shopping district where the regular stores are located. As before, this is a pattern that is also seen in the national sample.

By inspecting the data alone, it can reasonably be inferred that geographic market segmentation is not a driver of the outlet store strategy. The two store formats serve nearly identical locations, and often attract the same consumers. This leaves one last hypothesis

[1] CBSAs consist of metropolitan statistical areas and micropolitan areas–collectively areas based on urban centers of at least 10,000 people and economically relevant adjoining areas.

Figure 2: Consumers in Indianapolis, IN

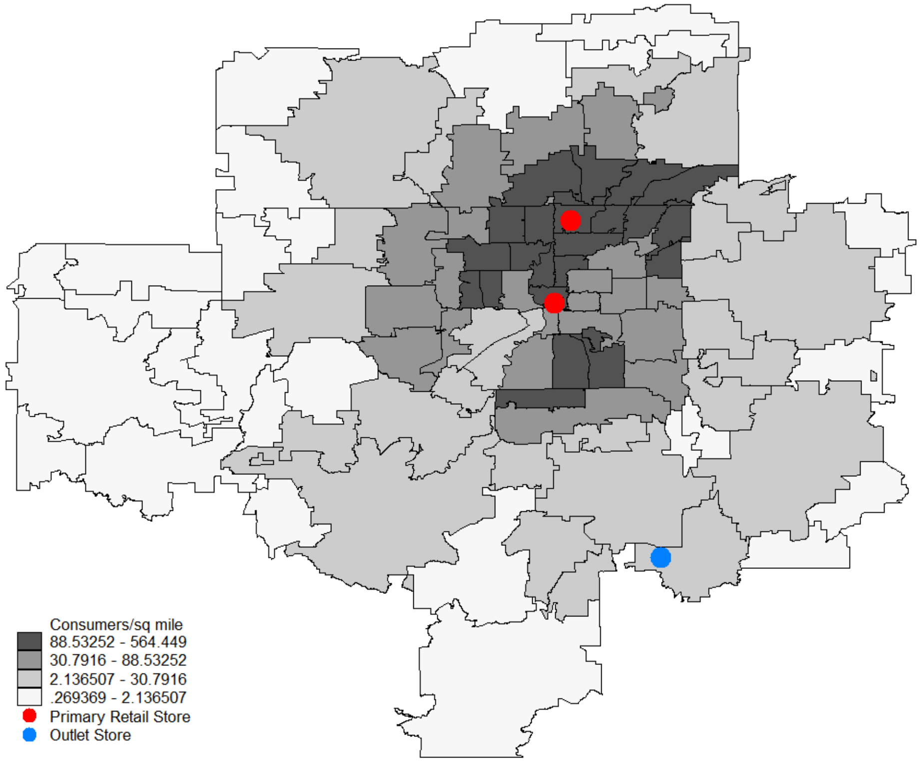
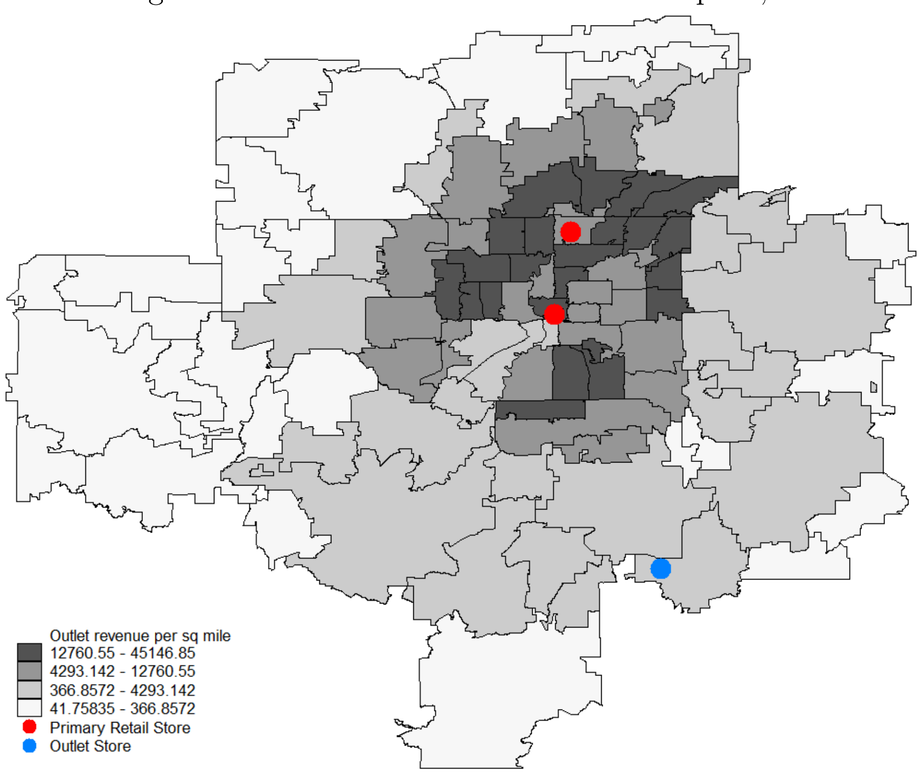


Figure 3: Outlet store revenues in Indianapolis, IN



to consider: that the firm's selling strategy is designed to implement price discrimination through consumer self-selection.[2]

4.3 Consumer self-selection

This paper focuses on illustrating how outlet stores induce a segment of consumers to travel for discounts. While, as Tables 3 and 4 show, consumers do not markedly differ in their observable attributes by format choice, this does not preclude them from differing in their preferences. In the following section, I lay out a demand model that permits heterogeneity in unobserved consumer tastes. Among other uses, estimation of the model's parameters will allow me to more fully characterize the differences in regular store and outlet store patrons. This step illustrates how the firm's selling strategy achieves a sorting of consumers according to their preferences.

5 Demand

In this section, I present a model of consumer demand, which takes on a nested mixed logit form. I proceed to discuss how I estimate the parameters of the model using transactions data from the firm. Finally, I present the results of demand estimation and discuss what they imply about the function of outlet stores as a tool for price discrimination.

Demand model. Since the typical consumer chooses between multiple locations, it is natural to think of her purchase decision as consisting of a store choice followed by a product choice. Conditional on her store choice, the indirect utility that a consumer i derives from purchasing product j in month t is

$$u_{ijt} = \xi_j - (\alpha + \zeta_i)p_{jt} - (\beta + \eta_i)age_{jt} + \epsilon_{ijt}. \tag{1}$$

That is, her utility is determined by: the intrinsic quality of the product, ξ_j; the product's price p_{jt} at time t; time that has elapsed since the product was introduced, denoted age_{jt}; and an idiosyncratic demand shock ϵ_{ijt}. I assume that consumers vary in their price sensitivity according to deviations ζ_i from the mean level α, and in their taste for new products according to deviations η_i from the mean level β. Utility from the outside good is normalized to $u_{i0t} = \epsilon_{i0t}$. I also assume that ϵ_{ijt} is i.i.d. type-I extreme value.

At the store, the consumer chooses the product that gives her the highest utility. Given the distributional assumption on ϵ_{ijt}, this implies that the expected utility consumer i derives from a store k's product assortment J_k is the inclusive value

[2] Here "geographic segmentation" is taken to be synonymous with third-degree price discrimination, and "self-selection" with second-degree price discrimination.

$$IV_{ik} = \log\left(\sum_{h\in J_k} \exp(\xi_h - (\alpha+\zeta_i)p_{ht} - (\beta+\eta_i)age_{ht})\right). \tag{2}$$

Consumers choose which store to visit based on store characteristics in addition to their expected utility from the available products. Consumer i's utility from visiting store k is

$$\tilde{u}_{ikt} = \tilde{\xi}_k + \lambda IV_{ik} - (\gamma + \nu_i)distance_{ik} + \tilde{\epsilon}_{ikt}. \tag{3}$$

A desirable feature of the data is that each consumer's billing zip code is observed, allowing for a focus on the role of travel distance in consumer choices. In addition, I allow for individual deviations ν_i from the mean level of sensitivity to travel γ. The parameter λ governs substitution patterns between products and stores by indicating the correlation in unobserved product characteristics within each store. The fixed effect $\tilde{\xi}_k$ captures the attractiveness of features of store k that are unrelated to the products within it or its distance from consumers. I normalize utility from no store visit to $\tilde{u}_{i0t} = \tilde{\epsilon}_{i0t}$ and again assume that $\tilde{\epsilon}_{ikt}$ is i.i.d. type-I extreme value.

These distributional assumptions imply that the probability that consumer i purchases product j in store k is

$$\begin{aligned} P_{it}(j_k) &= P_{it}(j|k)P_{it}(k) && (4)\\ &= \frac{\exp(\xi_j - (\alpha+\zeta_i)p_{jt} - (\beta+\eta_i)age_{jt})}{1+\sum_{h\in J_k}\exp(\xi_h - (\alpha+\zeta_i)p_{ht} - (\beta+\eta_i)age_{ht})} && (5)\\ &\quad\times\frac{\exp(\tilde{\xi}_k + \lambda IV_{ik} - (\gamma+\nu_i)distance_{ik})}{1+\sum_{l\in K_i}\exp(\tilde{\xi}_l + \lambda IV_{il} - (\gamma+\nu_i)distance_{il})} && (6)\end{aligned}$$

where K_i is the set of stores in consumer i's market.

I further assume that $\zeta_i \sim N(0,\sigma)$ and $[\ \eta_i \quad \nu_i\ ]' \sim N(0,\Sigma)$. This allows for an arbitrary correlation between sensitivity to travel and taste for new products. Correlations between these values and price sensitivity are restricted to zero. This restriction is partially relaxed in the appendix, which also contains a discussion of possible implications.

Note that, based on the specified model and the granularity of the data, consumers are identical up to their billing zip codes. Consequently, the predicted market share of product j in store k at the zip code z where consumer i resides is

$$s_{zt}(j_k) = \int_i P_{it}(j_k)df(\eta_i,\nu_i;\sigma,\Sigma), \tag{7}$$

where f is a multivariate normal pdf.

Let n_{zjkt} be the number of consumers in zip code z that purchase product j at store k in month t. The log-likelihood function given a set of parameter values and fixed effects $\Theta = (\alpha, \beta, \gamma, \lambda, \sigma, \Sigma, \{\xi_j\}, \{\tilde{\xi}_k\})$ is

$$l(\Theta) = \sum_t \sum_{k \in K_z} \sum_{j \in J_{kt}} \sum_z n_{zjkt} \log s_{zt}(j_k), \tag{8}$$

where K_z is the set of stores geographically accessible from zip code z.

Market sizes and outside options. For estimation purposes, the market size for each zip code is the total number of unique consumers who made a purchase within the entire sample. The assumption is that consumers who do not make any purchases within the 5-year period are not part of the market. If a consumer purchases a non-main category product from store k, then she is counted as visiting store k and choosing the outside option. If a consumer is not observed during a period, then she is counted as not having visited a store.

Identification. The firm's pricing practices allows for the consistent estimation of α amd σ without the use of instrumental variables techniques. To begin with, the firm implements a national pricing regime, thereby eliminating any systematic pricing differences between markets. Within-product variation in prices is generated by two sources. The first is randomly implemented store-wide promotions. These typically take the form of discounts that apply to all of the products in-store. The second is a systematic marking down of products over time. Table 6 shows through a projection of prices on product fixed effects, an outlet dummy, and age that most of the variation in prices is accounted for by the included variables, while the leftover variation falls within the scope of the randomized promotions.

Table 6: Pricing equation

Variable	**Coefficient**	**St Dev**
constant	5.33	0.15
outlet	-0.42	0.0029
log(age)	-0.43	0.0022
depvar	log(price)	
product FE	yes	
R2	0.8965	

The inclusion of product and store fixed effects in the estimation absorbs all unobserved quality differences between products and stores outside of age and distance. This also addresses potential endogeneity concerns with respect to the assignment of products to particular stores.

Table 7 outlines the result of the estimation procedure. All estimated coefficients have the expected sign: higher prices, older ages, and farther distances adversely affect utility. The Choleski decomposition of covariance matrix Σ is precisely estimated and implies a large variance in travel sensitivity and taste for new products. The estimates indicate a high correlation between travel sensitivity and taste for new products: consumers who highly dislike traveling also dislike buying old merchandise.

Table 7: Demand estimates

	coef	se
Product level		
price	-2.327	0.503
σ_{price}	0.344	0.109
age	-2.621	0.681
Store level		
IV	0.442	0.183
distance	-0.912	0.079
chol(Σ)		
(1,1)	0.908	0.297
(2,1)	0.435	0.178
(2,2)	0.209	0.172
Implied covariances		
σ_{age}	0.908	
σ_{dist}	0.483	
$\rho_{age,dist}$	0.629	
N	7,566,195	
l	1,832.09	
product fixed effects	yes	
store fixed effects	yes	

An interpretation of the estimated coefficients for price, age, and distance is that the average consumer would have to be compensated roughly \$100 in order to maintain her level of utility given a one-year increase in the age of a product or a 20-mile increase in travel distance. The λ estimate implies a moderate correlation in demand shocks within each store.

Market segmentation. Estimating the underlying parameters of consumer preferences allows for a description of consumers based on unobservable characteristics. Here I use the estimates to expound on the differences between consumers who buy goods from regular stores and those who buy goods from the outlet stores. I do this by using my demand model to predict purchase behavior given the available products for different types of consumers. Recall that consumers and their choices differ in multiple ways: (1) within each market, they vary by home zip code and thus perceive relative travel distances differently, (2) store availability and assortment differ between markets, and (3) consumers in all locations differ in their travel sensitivity and taste for new products.

Table 8 adds to the information in Table 3 through demand estimation. Whereas the data shows that consumers do not significantly differ by income and other purchase behavior depending on which format they choose, estimation reveals that they differ greatly in travel sensitivity and taste for new products.

Table 8: Market segmentation by consumer tastes

Consumer values ($) for:	Regular Stores	Outlet Stores
20-mile travel distance increase	71.83	36.21
	(12.47)	(11.08)
1-year product age increase	51.97	33.45
	(14.22)	(12.76)

This table lists consumer values in dollars for changes in store and product attributes. Standard errors are in parentheses.

The analysis in this section provides supportive evidence that through the firm's outlet store strategy, it segments consumers according to their underlying preferences for travel and newness. Discounts in outlet stores seem deep enough to cater to lower-value consumers, but not enough to cater to consumers who place a high premium on convenience and new arrivals.

A complete argument for these conclusions requires studying counterfactual store configurations and the associated consumer responses. The natural counterfactual scenario is one in which the firm chooses not to open locations in outlet malls. It would be insufficient, however, to simply remove these locations from the data and simulate purchase behavior. The firm would presumably charge different prices in its regular stores in the absence of outlet stores. Since outlet stores form an integral part of the firm's distribution strategy, removing them would also motivate changes in the how the firm stocks its regular stores.

The following section provides a framework for thinking about how the firm chooses prices and product assortments given its dual distribution strategy. The purpose of these models is to form a basis, together with the demand model, for predicting firm performance given a counterfactual distribution strategy.

6 Supply

In this section I develop a model of firm behavior with respect to price-setting and product assortment choice. This model permits a careful comparison of firm performance under counterfactual consumer characteristics and alternative distribution strategies, and hence sheds light on the profitability of outlet stores. This also allows an examination of the firm's costs, which serve as both a basis for the policy simulations and an indicator of the validity of the model's assumptions.

Two major assumptions are maintained throughout this section. The first is that the firm behaves like a monopolist, setting prices and product characteristics without strategic considerations. The second is that the firm's prices and product choices maximize profits conditional on store locations. I discuss each of these assumptions before describing the model.

The monopoly assumption is motivated by the firm's unique position in the industry. It has a 30-40 percent share of total industry revenues, and an even larger share in its particular psychographic segment. The next largest brand accounts for about 10 percent of industry revenues. Their products, however, retail at about the \$1,000 price point—much higher than our firm's average price of \$300. There is arguably little overlap between the market for our firm's products and the market for higher-end products such as those carried by the number two brand.[3]

The firm's dominant position also motivates the assumption that the firm is profit-maximizing. There may be very few firms for which this is a more appropriate assumption to make, given the firm's reputation not only in its category but also across industries. The firm consistently ranks among the top 10 firms across all industries in revenue per square foot of retail space, which is a standard performance metric among retailers.

I categorize firm decisions according to long- and short-term horizons. Long-term decisions concern store locations, stylistic product characteristics, and store capacities. Short-term decisions consist of pricing and the choice of product introduction rates. In my supply model, I take the firm's long-term decisions as exogenous, and treat the short-term decisions as endogenous.

I now proceed to describe the supply model in detail. First I discuss pricing. The monopoly pricing assumption, combined with the previous section's demand model, implies marginal costs for each product. I show how these marginal costs relate to observed product characteristics. Next I add endogenous product choice. The added features, combined with the pricing and demand models, pin down product development costs.

Prices. The firm sets prices in each period to maximize profit given store locations and product characteristics. The firm's profit function, conditional on product characteristics, is

$$\pi(\mathbf{p}_t) = \sum_{z,t} \left(M_z \sum_{k \in K_z} \sum_{j \in J_{kt}} s_{zt}(j_k)(p_{jt} - mc_{jt}) \right) \tag{9}$$

That is, per-product (j) profit in each zip code z and month t is price p_{jt} minus marginal cost mc_j times quantity sold $M_z s_{zt}(j_k)$, where M_z is market size and $s_{zt}(j_k)$ is market share as determined by Equation 7. Total profit is the sum over all products, periods, and geographical markets, where the set of products in each store is J_{kt} and the set of stores assigned to each zip code is K_z. Profit-maximizing prices satisfy the first-order conditions

$$\frac{d\pi}{dp_{ht}} = \sum_{z,t} M_z \sum_{k|h \in J_{kt}} \left(s_{zt}(h_k) + \sum_{j \in J_{kt}} \frac{\partial s_{zt}(j_k)}{\partial p_{ht}} (p_{jt} - mc_{jt}) \right) = 0 \tag{10}$$

for each product h and month t. Rewriting the conditions for each period (and suppressing time subscripts) as $\mathbf{s} + \Delta(\mathbf{p} - \mathbf{mc}) = 0$ where $\mathbf{s}_j = \sum_z \sum_{k|j \in J_k} M_z s_z(j_k)$, $\Delta_{h,j} =$

[3] There is little publicly available information with more precise figures—these market shares were relayed by the firm's executives. They also agree with the notion that competitors' pricing trends have little or no impact on the firm's pricing decisions.

$\sum_z \sum_{k|j \in J_{kt}} \frac{\partial s_z(j_k)}{\partial p_h}$, and $\mathbf{p}_j = p_j$, the marginal cost of each product in a given period is exactly identified using estimated demand coefficients:

$$\mathbf{mc} = \mathbf{p} + \Delta^{-1}\mathbf{s}. \tag{11}$$

I use Equation 11 to compute marginal costs for each product. Recall that observed prices are "contaminated" by randomized promotions, which conceivably cause departures from strict profit maximization. I make the operational assumption that the estimated parameters in Table 6 are profit-maximizing choices by the firm. I use the pricing equation from Table 6 to find predicted prices for each product, which I interpret as the fully endogenous component of prices that adheres to profit maximization. These are the prices I use to compute marginal costs for each period.

For descriptive purposes I project these marginal costs linearly onto product characteristics, and present the coefficients in Table 9.[4] The implied average marginal cost over all products closely resembles figures from industry reports and suggestions from the firm's executives. The estimated relationships between characteristics and marginal cost are also sensible: premium material costs more than basic, and larger silhouettes cost more to manufacture than smaller ones. This provides an indication of the validity of the pricing equation.

Table 9: Marginal Cost Estimates (in Dollars)

Characteristic	Coefficient	SE
constant	41.64	12.08
premium material	22.98	10.70
basic material	-5.23	0.01
silhouette 1	14.09	9.56
silhouette 2	2.00	5.76
silhouette 3	-29.81	16.83
silhouette 4	-13.12	6.34
silhouette 5	-4.06	2.84
silhouette 6	-7.34	0.56
silhouette 7	25.14	8.95
silhouette 8	24.66	2.43
OLS depvar		mcost
N		848
R-Squared		0.45

Product choice. The overall product design process is exceptionally complex for firms that produce fashion goods. There is an expansive number of dimensions to determine for each of a huge number of products to generate periodically. It is unfeasible to model product choice as it applies to every individual good. This necessitates a means of drastically reducing the

[4]These marginal costs are computed for the last period in the sample.

number of choice variables for the firm while focusing on the most relevant decisions to the research question.

An important dimension of product choice for the firm that is salient to studying the outlet store strategy is that of product lifespans in each format. By lifespan, I mean the amount of time a product is available for purchase in each format. Figure 1 shows how product lifespans are determined by the flow of inventory into, between, and out of store formats. New products flow into both formats when "original" and "factory" products are born (see Table 2). All products in the regular store are eventually transferred to the outlet store, where the last units of each style is sold.

One advantage of using the current dataset to study firm product choice is that the outlet store strategy provides a structure that delimits the firm's choice set. The technology that the firm uses to create product age-distance combinations—physically transferring products between formats—is completely transparent and can mostly be considered cost-neutral. This is in contrast to most other cases, where both product assembly technologies and cost structures are more complex.

Although the number of new products in each format can conceivably be modeled using existing techniques, the selection of which products to transfer or discontinue presents a different challenge. Because the firm offers such a large number of products, an attractive option is to think of the firm as targeting a joint probability of product characteristics rather than individual product attributes. A primary contribution of this paper is a demonstration of this novel approach to modeling multidimensional product differentiation.

Specifically, I assume that store locations and capacities are given. Let C_k be the number of items that store k can display on its shelves. I assume that in each period, each store k of format $fmt \in \{regular, outlet\}$ takes C_k draws from the corresponding master set of products, described by the distribution of product characteristics ϕ_{fmt}. Let $\phi_{fmt} = f_{fmt} \times g_{fmt}$, where f_{fmt} is the distribution of endogenous product characteristics (product ages in this application) and g_{fmt} governs the exogenous characteristics (summarized here by ξ_j).[5] The firm's objective is to choose the profit-maximizing shapes of $f_{regular}$ and f_{outlet}.

In order to make this problem tractable, I propose to construct f_{fmt} using a set of parametric distributions. Industry logistics and the data suggest a natural choice for these distributions and a direct interpretation of their parameters. Consider these assumptions on product assortment:

1. Original products in the regular format have an average probability x of being transferred to the outlet format in the next period
2. Factory products in the outlet format have an average probability y of being retired in the next period
3. Original products in the outlet format have an average probability z of being retired in the next period

[5] Treating ξ_j as exogenous can be rationalized by the fact that the firm usually cannot ascertain the appeal of a product to consumers before it is taken to market.

Figure 4: Empirical versus simulated product age densities in regular format

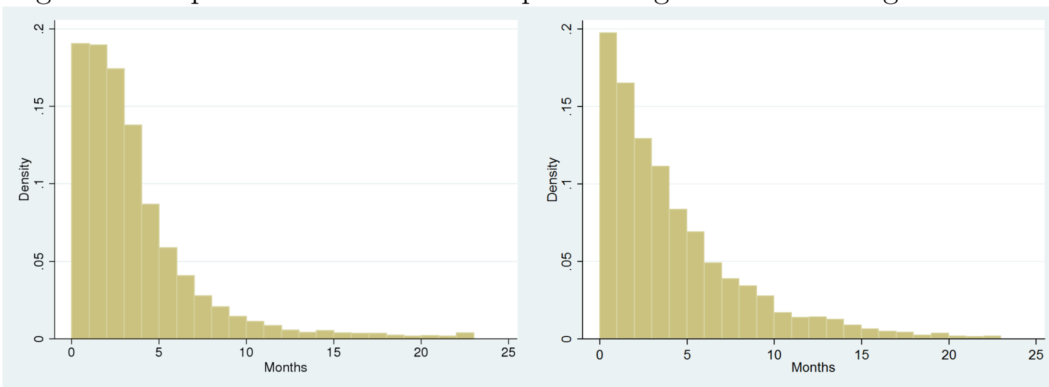
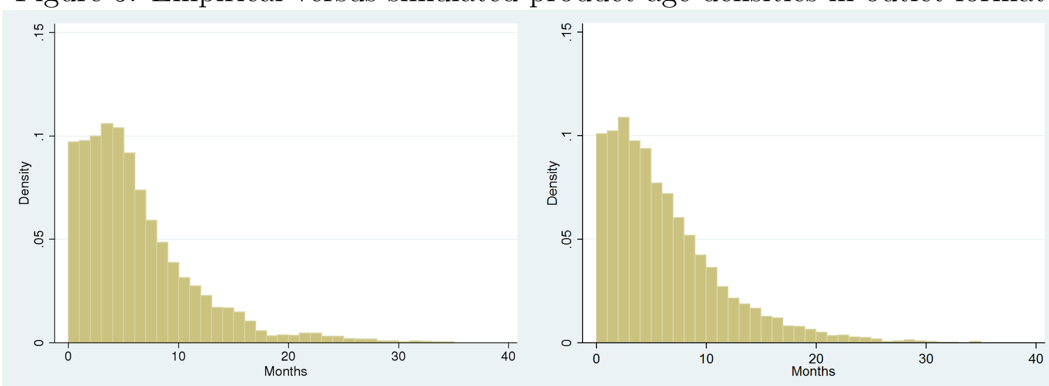


Figure 5: Empirical versus simulated product age densities in outlet format



4. Factory goods make up a proportion α of goods in the outlet format

These assumptions imply that if X is product age in the regular format and Y is product age in the outlet format then

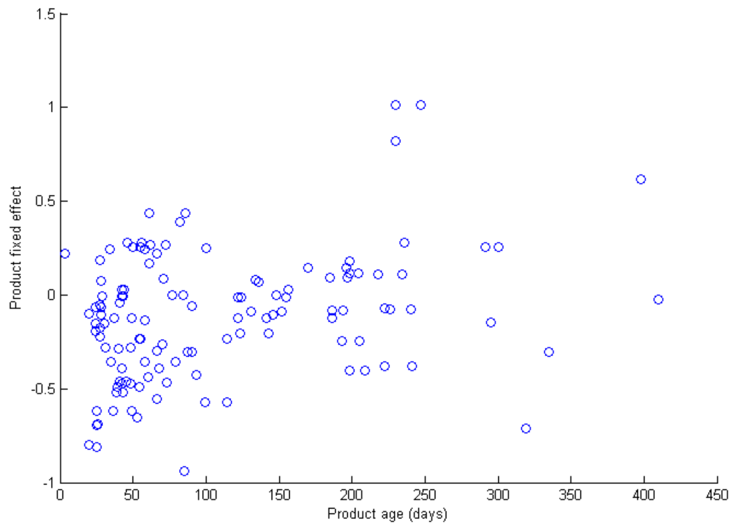
$$X \sim Geometric(x) \tag{12}$$

$$Y = \begin{cases} W & \text{with probability } \alpha \\ X + Z & \text{with probability } 1 - \alpha \end{cases} \tag{13}$$

where $W \sim Geometric(y)$ and $Z \sim Geometric(z)$

By adjusting the stopping probabilities x, y, and z, the firm can control the relative distributions of product age in each store format. These probabilities also pin down the portion of products that are new introductions in each period: the share of original products that are newly introduced in a period is simply x and the share of new factory products is y. Figures 4 and 5 illustrate how closely this parameterization resembles the observed distribution of product characteristics. The simulated densities (right) are generated using moments of the empirical densities (left).

Figure 6: Better products are longer-lived in regular stores



Original products in the regular format are not transferred to outlets at random. Products that perform better in sales, and thus are presumably of higher quality, have longer lifespans in regular stores. Figure 6 plots ξ_j against age_{jt} for the regular store offerings in the Indianapolis example from Figure 2. In the language of the exposition above, the distribution of endogenous characteristics $f_{regular}$ is dependent on that of exogenous characteristics $g_{regular}$. I keep the form of this dependence fixed by allowing the firm to adjust the speed of product turnover but not the order by which products are transferred according to their ξ_j.

Adjusting the restocking probabilities, and consequently the rate of new product introduction, has implications on per-period costs. Assuming that the firm chooses to maintain a fixed number of products in its universal offer set (i.e. the set from which store k draws C_k products), the cost per period $C(x, y)$ of implementing a given age distribution must depend on the number of new product introductions it requires. I use a linear function

$$C(x, y) = ax + by \tag{14}$$

to represent these costs.

I assume that the firm chooses product choice parameters x, y, z, and α once and prices $\mathbf{p}_t$ every period to maximize expected profit

$$E(\pi|x, y, z, \alpha, \mathbf{p}_t) = \sum_{fmt} \int \sum_{z,t} M_z \sum_{k} \sum_{j \in J_{kt}} s_{zt}(j_k)(p_{jt} - mc_{jt}) df_{fmt}(x, y, z, \alpha) - C(x, y) \tag{15}$$

This allows me to identify cost parameters a and b exactly through the first order conditions of profit maximization: $\partial E(\pi)/\partial x = \partial E(\pi)/\partial y = 0$. I solve these equations numerically for a and b, and present the implied product development costs in Table 10.

Before proceeding to discuss the fixed cost solutions, I describe how I compute the expected profit for perturbations around the observed x and y. First I sort the C_k products according to age within each store k. This allows me to fix the dependence of the stocking priorities on ξ_j. Given stocking probabilities x and y, I make ns sets of C_k draws from the distributions specified in Equations 12 and 13.[6] I replace the ages in the data with these draws, keeping the original order according to age constant. I then compute the average over corresponding profits for each of the ns draws.

Table 10: Implied product development costs

Product class	Parameter Value	Average stock	Fixed cost per unit
Original (a)	7,779,203	151	51,518
Factory (b)	19,034,202	433	43,959

The parameter values in Table 10 indicate the cost of replacing the entire stock of products, i.e., when $x = 1$ or $y = 1$. Dividing these values by the average stock of each class of product gives the fixed costs associated with developing each unit. I find that producing each style of product carries a fixed cost of about \$50,000, and that the fixed cost of producing an original product is significantly higher than the fixed cost of a factory product.

With the model of price-setting and product introduction discussed in this section, together with the fixed and marginal costs that they imply, counterfactual store configurations can now be properly evaluated.

7 Policy Simulations

The basic question that this paper addresses is: Why do outlet stores exist? In this section, I answer this question by simulating situations in which the firm pursues selling strategies that exclude outlet store retail. For each of these policy simulations, I use the supply-side model in Section 6 to specify how the firm would change its pricing and product introduction rates in response to changes in store locations. The demand model from Section 5 then shows how consumers would react to these changes in firm strategy. I find that outlet stores serve to expand the firm's market to include consumers who are more sensitive to prices, less averse to travel, and less particular about product ages. Furthermore, the assortment in outlet stores is chosen to prevent higher-value consumers from preferring to visit outlet stores over regular stores.

Test market. I use a representative market over which to perform policy simulations, in order to clearly demonstrate the effects of each experiment. I then show that the same conclusions are reached from running these experiments over the national sample.[7] The test market is the Indianapolis-Carmel Metropolitan Statistical Area in July 2007, maps of which were presented in Section 2. This market is representative of the firm's markets both

[6] $ns = 50$ in this version of the paper.

[7] This section is in progress.

in terms of the demand profile and the firm's store and product configurations. Table 11 lists store attributes and some performance measures in this market.

Table 11: Test market store characteristics

Store:	Regular 1	Regular 2	Outlet
Number of products	60	72	165
Average price	313.46	329.39	154.28
Average product age (mo)	13.14	13.49	20.04
Average distance (mi)	11.34	9.39	30.60
Units sold	148	217	967
Revenue	29,861.96	50,083.60	119,057.12

7.1 No outlet stores

The most natural policy experiment to run involves simply removing the outlet stores. Many large retail firms choose not to operate outlet stores. Although a careful comparison between firms is hard to make, it can be argued that these firms' selling strategies are similar to the firm's regular store strategy taken alone in several respects. For instance, the firm's regular stores are of similar size, configuration, and location to those of Louis Vuitton, even though there are no Louis Vuitton outlet stores.

Table 12 contains the results of this counterfactual as they pertain to the supply-side responses. Column 1 contains the actual average prices, product ages, and revenues, which are used as a baseline. Column 2 shows that revenues in regular stores increase when the outlet store is closed, even when prices and assortment in the regular stores remain the same. Column 3 shows that the firm would lower prices in regular stores in the absence of outlet stores, even if it could not change the assortment (see Appendix B for details on finding optimal prices). Column 4 shows that the firm would choose to make fewer product introductions if outlet stores did not exist, resulting in an increase in average product age in these stores.

The story is rounded out by looking at details of the demand-side response, which are listed in Table 13. Closing the outlet store initially results in a very small increase in regular store revenues because few of the consumers who shopped at the outlet store switch to regular stores. When allowed to change product characteristics, the firm lowers quality and price in the regular stores to cater to the lower-value consumers. However, even given this flexibility, the firm is unable to serve the full range of consumers that it can with the outlet stores present.[8]

[8] A clearer picture of which consumers are served can be presented through heat maps in consumer taste space, to follow.

Table 12: No outlet stores (supply response)

	1	2	3	4
Regular 1				
Price	313.46	313.46	280.21	220.73
Age	13.14	13.14	13.14	15.12
Revenue	29,862	32,771	35,911	36,125
Regular 2				
Price	329.39	329.39	302.51	250.03
Age	13.49	13.49	13.49	15.12
Revenue	50,084	55,831	62,200	67,830
Outlet				
Price	154.28	-	-	-
Age	20.04	-	-	-
Revenue	119,057	-	-	-
Total revenue	199,003	88,602	98,111	103,955
Variable profit	106,728	62,042	71,389	73,518

Prices and product ages are averages over each store.
Columns indicate:
1 - Baseline
2 - Outlet store closed
3 - Prices reoptimized
4 - Prices and product ages reoptimized

Table 13: No outlet stores (demand response)

	1	2	3	4
Regular 1				
Distance aversion	73.98	70.04	65.32	63.42
Age aversion	45.17	42.99	39.13	35.16
Regular 2				
Distance aversion	81.09	80.82	72.15	70.21
Age aversion	42.65	41.53	37.26	32.88
Outlet				
Distance aversion	32.55	-	-	-
Age aversion	25.90	-	-	-

Consumer values are averages over each store. Distance aversion is the dollar equivalent to a consumer of a 20-mile increase in travel distance. Age aversion is equivalent to a 1-year increase in product age.
Columns indicate:
1 - Baseline
2 - Outlet store closed
3 - Prices reoptimized
4 - Prices and product ages reoptimized

7.2 Random assortment

The assignment of products to either regular stores or outlet stores forms an important part of the firm's selling strategy. In this subsection, I show the value of the firm's observed assortment strategy by comparing its observed performance with that achieved by a counterfactual assortment strategy in which products are randomly assigned to stores. This random assignment results in a configuration in which regular and outlet stores contain roughly identical assortments.[9,10] This counterfactual strategy resembles that of firms that open stores in outlet malls, but do not distinguish the assortment in these stores from those in their non-outlet locations.

Table 14 describes the resulting average product characteristics in these stores. Here I allow the firm to adjust prices, so that in both cases prices are profit-maximizing conditional on product assortments. Jumbling the products results in near-identical average product qualities between stores, but prices are still much lower in the outlet store. This suggests that the bulk of discounting in outlet stores is to compensate for the inconvenience associated with longer travel times.

Table 14: Randomized product distribution–supply

Assortment:	Actual	Randomized
Regular 1		
Age	13.14	16.92
Price	313.46	300.12
Regular 2		
Age	13.49	17.43
Price	329.39	302.18
Outlet		
Age	20.04	17.18
Price	154.28	170.48

As reported in Table 15, the firm's performance suffers under a random assignment of products to stores. Revenues in all stores decrease, and consumers are less different between formats. This should be unsurprising, given that the products are less different between formats. My hypothesis is that sorting works exceptionally well because there is a positive correlation between consumer travel sensitivity and taste for newness. To test this hypothesis, I run the same counterfactual but under a supposed form of consumer heterogeneity in which there is zero correlation between travel sensitivity and tastes for newness.

Table 16 has the results of this experiment. As anticipated, randomizing assortment has less of an effect when consumer tastes for the two attributes are uncorrelated. There was little sorting to begin with, so the decrease does not come with very big a cost.

[9] Recall that a product is unique only up to its fixed effect ξ_j, its price p_j, and its vintage age_j.
[10] Outlet stores will still have more shelf space than regular stores.

Table 15: Randomized product distribution (actual tastes)

Assortment:	Actual	Randomized
Regular 1		
Distance aversion	73.98	70.87
Age aversion	45.17	41.74
Regular 2		
Distance aversion	81.09	75.42
Age aversion	42.65	39.21
Outlet		
Distance aversion	32.55	41.53
Age aversion	25.90	35.23
Total revenue	199,003	173,374
Variable profit	106,728	85,150

Table 16: Randomized product distribution (uncorrelated tastes)

Assortment:	Actual	Randomized
Regular 1		
Distance aversion	54.82	53.88
Age aversion	40.84	39.15
Regular 2		
Distance aversion	62.46	60.32
Age aversion	37.28	36.71
Outlet		
Distance aversion	41.38	43.28
Age aversion	31.84	33.18
Total revenue	153,432	151,883
Variable profit	73,648	71,832

7.3 Centrally-located outlet stores

It seems plain to see that the firm pursues a 'damaged goods' strategy by selling a portion of its goods in distant locations. In order to confirm this hypothesis, I run a third set of counterfactuals in which outlet stores are moved to central locations. I show that (i) revenues decrease, (ii) the firm would make fewer product introductions in the outlet format, and (iii) cater to a narrower range of consumers. I also show that the benefit of damaging goods in this fashion is increasing in the taste correlation, in the same sense as in Section 7.2.

An alternative explanation to damaged goods is that firms locate in outlet malls to take advantage of lower rents. Outlet malls on average set a monthly rent of \$29.76 per square foot, which can be dwarfed by rents in the most prestigious retail locations (Humphers 2012). However, this rent is close to the average for retail space in many urban centers—implying that the firm could choose to costlessly relocate its outlet stores closer to its target market.

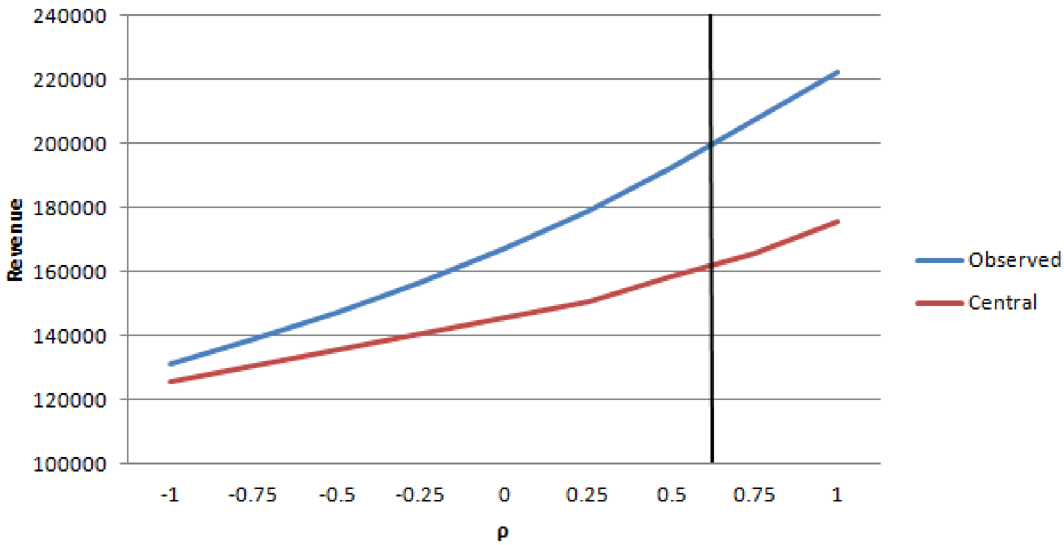
Table 17 presents the supply-side results of the experiment in which the outlet store is moved into the central shopping district. Notably, while prices are less variable now (regular store products are cheaper and outlet store products are more expensive), quality along the age dimension is more variable (regular store products are slightly newer and outlet store products are much older). Denied the ability to differentiate products according to location, the firm increases the level of differentiation according to age. The range of consumers that the firm is able to reach, nevertheless, is similar to the case in which the outlet store is simply shut down.

Table 17: Outlet moved to center (supply response)

Outlet location:	Actual	Central
Regular 1		
Age	13.14	12.78
Price	313.46	308.21
Distance	11.34	11.34
Regular 2		
Age	13.49	13.00
Price	329.39	314.77
Distance	9.39	9.39
Outlet		
Age	20.04	23.89
Price	154.28	204.76
Distance	30.60	10.69
Total revenue	199,003	162,601
Variable profit	106,728	70,419

Figure 7 shows the relationship between the relative profitability of outlet store retail and the correlation between consumer tastes for quality and convenience. It is additional evidence for the idea that the firm exploits the correlation in consumer attributes through its outlet store

Figure 7: Central outlet policy vs taste correlation



strategy. It also suggests a plausible reason for why outlet stores have become so popular among clothing and fashion firms, but not so much in other industries: tastes for quality and convenience may not be so strongly correlated elsewhere.

These counterfactuals show that adding outlet stores helps the firm in many ways. First, it extends the firm's market to include consumers who are not averse to traveling and less desirous of new things. Since these are the same people in the data, it makes sense for the firm to populate its outlet stores with older products. This has the additional benefit of making outlet store products less attractive to higher-value consumers, thus preventing cannibalization.

8 Conclusion

Owning and operating outlet stores constitutes a major component of many firms' distribution strategies, particularly in the clothing and fashion industries. It is an interesting practice that continues to evolve and gain popularity. Yet there has been little written in the marketing and economics literatures that speaks to the reasons for the success of outlet stores, or the mechanisms by which they improve firm performance. The availability of new sales data from a major fashion goods manufacturer and retailer offers a unique opportunity to empirically investigate how outlet stores work.

This paper shows that outlet stores provide several benefits as a tool of price discrimination. Outlet stores allow the firm to serve lower-value consumers without lowering prices faced by its regular store clientèle. By stocking outlet stores with less desirable products, the firm exploits the positive correlation between consumers' travel sensitivity and taste for quality. Prices are low in outlet stores, but not low enough to attract consumers who value quality and convenience the most.

The model of product choice in this paper suggests a benefit of running outlet stores apart from its price discrimination uses: it allows the firm to make more frequent new product introductions in its regular format. The firm offers more new products every period in its regular stores both to increase the attractiveness of regular store offerings relative to those in outlet stores and because it stocks outlet stores with older, less attractive products from the regular stores. This can conceivably counter the threat that is most associated with outlet stores: that it results in the dilution of prestige brands. Outlet stores may actually enable the firm to improve its regular store products, which typically form the basis of a fashion brand's image.

Lessons from outlet store retail have wide applicability to questions of product line design and price discrimination. Outlet stores are a specific response to the apparent heterogeneity in tastes for quality and convenience among fashion shoppers. Similar responses by firms to consumer tastes can be observed in the electronics and travel industries. The notion that the correlation of characteristics in a firm's product space ought to resemble the correlation of consumer tastes for them may be useful to many firms.

The key insight is that multidimensionality in consumer preference heterogeneity matters for product line design. Firms that seek to optimize their product offerings must take into account how tastes vary for different product characteristics, and what the correlations between those tastes are. This is not a new discovery: the extant theoretical literature on multidimensional screening emphasizes the sensitivity of the optimal allocation set by the principal to the agent's value correlations. This is, however, the first demonstration of its importance in an actual business setting. The choice of whether to operate outlet stores hinges on a market landscape in which consumers who are most willing to travel to outlet malls value quality the least.

There are many possible directions for future research. Outlet stores constitute a single aspect of a consolidated selling strategy that has become standard among fashion goods firms. Other parts of this strategy include price skimming, targeted coupons, and loyalty programs. Many of these components operate on the intertemporal dimension of durable goods demand. It would be interesting to see how they build on the firm's overall product lining strategy by adding yet additional dimensions.

This setting is also a prime vehicle for exploring alternative theories of consumer behavior with respect to fashion goods demand. Consumers, for instance, may conceivably choose products based on their capacity to signal status. Meanwhile, fashion goods may have characteristics that are discernible to some consumers but not others (for instance, whether a product is sold exclusively in outlet stores). This presents a unique product design challenge to firms that wish to exploit these consumer preferences.

References

[1] Armstrong, Mark. "Multiproduct nonlinear pricing." Econometrica: Journal of the Econometric Society (1996): 51-75.

[2] Armstrong, Mark and Rochet, Jean-Charles. 1999. "Multi-dimensional Screening: A User's Guide. European Economic Review 43, 959-979.

[3] Bayus, Barry L., and William P. Putsis. 1999. "Product proliferation: An empirical analysis of product line determinants and market outcomes." Marketing Science 18.2: 137-153.

[4] Chen, Yongmin and Riordan, Michael H. 2012. "Profitability of Product Bundling." Mimeo, Columbia University.

[5] Coughlan, Anne T. and Soberman, David. 2005. "Strategic Segmentation Using Outlet Malls." International Journal of Research in Marketing 22, 61-86.

[6] Crawford, Gregory. 2012. "Endogenous Product Choice: A Progress Report." International Journal of Industrial Organization 30, 315-320.

[7] Crawford, Gregory S., Oleksandr Shcherbakov, and Matthew Shum. "The welfare effects of monopoly quality choice: Evidence from cable television markets." Unpublished paper (2011).

[8] Deneckere, Raymond J., and R. Preston McAfee. "Damaged goods." Journal of Economics & Management Strategy 5.2 (1996): 149-174.

[9] Desai, Preyas S. "Quality segmentation in spatial markets: When does cannibalization affect product line design?." Marketing Science 20.3 (2001): 265-283.

[10] Dobson, Gregory, and Shlomo Kalish. "Positioning and pricing a product line." Marketing Science 7.2 (1988): 107-125.

[11] Draganska, Michaela, and Dipak C. Jain. "Consumer preferences and product-line pricing strategies: An empirical analysis." Marketing Science 25.2 (2006): 164-174.

[12] Draganska, Michaela, Michael Mazzeo, and Katja Seim. "Beyond plain vanilla: Modeling joint product assortment and pricing decisions." QME 7.2 (2009): 105-146.

[13] Fan, Ying. "Ownership consolidation and product quality: A study of the us daily newspaper market." Working Paper, University of Michigan, 2010.

[14] Green, Paul E., and Abba M. Krieger. 1985. "Models and heuristics for product line selection." Marketing Science 4.1: 1-19.

[15] Heiss, Florian, and Viktor Winschel. "Likelihood approximation by numerical integration on sparse grids." Journal of Econometrics 144.1 (2008): 62-80.

[16] Humphers, Linda. 2012. "State of the Outlet Industry." Value Retail News.

[17] Johnson, Justin P. and Myatt, David P. 2006. "On the Simple Economics of Advertising, Marketing, and Product Design." American Economic Review 96(3), 756-784.

[18] Kekre, Sunder, and Kannan Srinivasan. 1990. "Broader Product Line: A Necessity To Achieve Success?" Management Science 36(10), 1216-1232.

[19] Kuksov, Dmitri. "Buyer search costs and endogenous product design." Marketing Science 23.4 (2004): 490-499.

[20] Leslie, Paul. 2004. "Price Discrimination in Broadway Theater." The RAND Journal of Economics 35(3), 520-541.

[21] McBride, Richard D., and Fred S. Zufryden. "An integer programming approach to the optimal product line selection problem." Marketing Science 7.2 (1988): 126-140.

[22] McManus, Brian. 2007. "Nonlinear Pricing in an Oligopoly Market: The Case of Specialty Coffee." The RAND Journal of Economics 38(2), 512-532.

[23] Moorthy, K. Sridhar. 1984. "Market segmentation, self-selection, and product line design." Marketing Science 3.4: 288-307.

[24] Morrow, W. Ross, and Steven J. Skerlos. "Fixed-Point Approaches to Computing Bertrand-Nash Equilibrium Prices Under Mixed-Logit Demand." Operations research 59.2 (2011): 328-345.

[25] Mussa, Michael and Rosen, Sherwin. 1978. "Monopoly and Product Quality." Journal of Economic Theory 18, 301-317.

[26] Netessine, Serguei, and Terry A. Taylor. "Product line design and production technology." Marketing Science 26.1 (2007): 101-117.

[27] Qian, Yi, Eric Anderson, and Duncan Simester. Multichannel Spillovers from a Factory Store. No. w19176. National Bureau of Economic Research, 2013.

[28] Randall, Taylor, Karl Ulrich, and David Reibstein. "Brand equity and vertical product line extent." Marketing science 17.4 (1998): 356-379.

[29] Reibstein, David J., and Hubert Gatignon. "Optimal product line pricing: The influence of elasticities and cross-elasticities." Journal of Marketing Research (1984): 259-267.

[30] Rochet, Jean-Charles and Stole, Lars. 2003. "The Economics of Multidimensional Screening." Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress 1, 150-197.

[31] Schmalensee, Richard. 1984. "Gaussian Demand and Commodity Bundling." The Journal of Business 57(1), S211-S230.

[32] Veiga, Andre and Weyl, E. Glen. 2012. "Multidimensional Product Design." Mimeo, SSRN.

Appendix

A Estimation

Maximization of the log-likelihood function in Equation 8 is performed numerically in Matlab using Knitro's interior-point direct algorithm. Efficient integration over $f(\eta_i, \nu_i; \Sigma)$ of each share $s_z(j_k)$ in Equation 7 is achieved by quadrature on sparse grids (Heiss and Winschel, 2007). For each guess of Σ, its Choleski decomposition $chol(\Sigma)$ is taken and multiplied by the matrix of uncorrelated nodes to generate nodes with the corresponding covariance structure.

B Finding optimal prices

The policy simulations in Section 7 involve generating optimal prices given counterfactual product characteristics. This presents a nontrivial computational task given the sheer number of products and markets. Morrow and Skerlos (2011) provide a fixed-point approach to finding these prices that dramatically reduces the computational burden. In particular, they demonstrate that iteration on a rearrangement of Equation 11,

$$\mathbf{p} = \mathbf{mc} - \boldsymbol{\Delta}^{-1}\mathbf{s}$$

results in convergence to profit-maximizing prices for mixed-logit demand systems.

C Alternative covariance specifications

Section 5 describes a demand model that features consumer heterogeneity in coefficients for price, distance, and product age. The covariance in coefficients for distance and product age is estimated, but covariances between coefficients for price and distance, as well as price and product age, are set to zero. Future versions of this model will allow for nonzero covariances between all random coefficients. In this section, I estimate pairwise covariances between coefficients for price, distance, and product age, and discuss possible implications on the policy simulations.

I maximize the same likelihood as in Equation 8 but with alternative restrictions on the joint covariance matrix of ζ_i, η_i, and ν_i—random shocks to coefficients for price, travel distance, and product age, respectively. Table A presents estimates for the case in which $\eta_i \sim N(0, \sigma)$ and $[\ \zeta_i \quad \nu_i\ ]' \sim N(0, \Sigma)$, and Table B presents estimates for that in which $\nu_i \sim N(0, \sigma)$ and $[\ \zeta_i \quad \eta_i\ ]' \sim N(0, \Sigma)$ While the estimated $\rho_{price,age}$ and $\rho_{price,dist}$ are significant, they are much smaller in magnitude than the estimate of $\rho_{age,dist}$. This suggests that there is less scope for the study of counterfactuals in which these correlations are made closer to zero.

Future work will involve estimating the full covariance matrix, i.e. one in which each element of Σ is unconstrained, where $[\ \zeta_i \quad \eta_i \quad \nu_i\ ]' \sim N(0, \Sigma)$. Substantial variation in prices, product ages, and store locations relative to consumers, both within and across products, allows

Table A: Nonzero covariance in price and distance coefficients

	coef	se
Product level		
price	-2.199	0.690
age	-2.863	0.523
σ_{age}	0.207	0.183
Store level		
IV	0.434	0.181
distance	-1.148	0.182
chol(Σ)		
(1,1)	0.095	0.023
(2,1)	-0.077	0.012
(2,2)	0.511	0.364
Implied covariances		
σ_{price}	0.308	
σ_{dist}	0.517	
$\rho_{price,dist}$	-0.148	
N	7,566,195	
l	1,831.78	
product fixed effects	yes	
store fixed effects	yes	

for the full identification of Σ. The computational burden of estimating two additional random coefficients, however, can be much greater given both the base size of the data and the additional (Gauss-Hermite quadrature) sample points needed to estimate a 3×3 covariance matrix for a given level of accuracy. At the accuracy in the current estimation, 53 weighted sample points are used to simulate the bivariate normal distribution. To simulate a trivariate normal at the same level of accuracy requires 165 weighted sample points, effectively tripling the number of simulated observations (Heiss and Winschel 2008).

While the pairwise correlations estimated in this section imply more limited interaction between demand coefficients for price and product characteristics than between product age and store convenience, allowing for all of these dependencies simultaneously in estimation may either temper or accentuate the estimated parameter of interest in Chapter **??**: the correlation in tastes for product age and store convenience. They may also affect the magnitudes of the predicted changes in performance in Chapter **??**'s counterfactuals. They would not, however, alter qualitative results about how the profitability of multidimensional product differentiation depends on the correlation in consumer tastes for each quality dimension.

Table B: Nonzero covariance in price and product age coefficients

	coef	se
Product level		
price	-2.468	0.722
age	-3.215	0.838
Store level		
IV	0.410	0.262
distance	-0.839	0.096
σ_{dist}	0.472	0.086
chol(Σ)		
(1,1)	0.374	0.148
(2,1)	-0.084	0.024
(2,2)	0.268	0.096
Implied covariances		
σ_{price}	0.374	
σ_{age}	0.281	
$\rho_{price,age}$	-0.299	
N	7,566,195	
l	1,830.92	
product fixed effects	yes	
store fixed effects	yes	