

Retail Competition and the Dynamics of Consumer Demand for Tied Goods*

Wesley R. Hartmann

Phone: 650-725-2311, Email: hartmann_wesley@gsb.stanford.edu

Harikesh S. Nair

Phone: 650-736-4256, Email: harikesh.nair@stanford.edu

Stanford Graduate School of Business
Stanford University, 518 Memorial Way, Stanford, CA 94305-5015

First version: July 19, 2007

This version: October 2, 2008

Abstract

We present a demand system for tied goods sold through competing retail formats. The demand system formally incorporates dynamics arising from the tied-nature of the products, as well as those arising from stockpiling considerations induced by the storable and durable nature of the products sold. The model endogenizes the retail format at which consumers choose to stockpile inventory, thereby facilitating measurement of long-run retail substitution effects. The model also yields estimates of complementarities within, and substitution across, competing systems of tied-goods. We present an empirical application to an archetypal tied-goods category, razors and blades. Implications of measured effects for manufacturer pricing when selling the tied products through an oligopolistic downstream retail channel are discussed.

Keywords: tied goods, retail competition, dynamic discrete choice, long-run effects, vertical channels, razor-blade market.

* We thank the review team, Michaela Draganska, Jean-Pierre Dubé, Ganesh Iyer, Jim Lattin, Philip Leslie, Yesim Orhun, Peter Reiss, K. Sudhir, Andrew Sweeting, Miguel Villas-Boas as well as seminar participants at the Marketing Science (2007), SICS (2007), CMRC-Yale (2008), and IIOC (2008) conferences, and at seminars at Cornell, Rochester, Santa Clara, Stanford, UT Austin and Yale SOM for their helpful feedback. We also thank Bart Bronnenberg, Pradeep Chintagunta and Carl Mela for their comments and help in obtaining the data used in this study. Ping Li and Madhavan Ramanujam provided exemplary research assistance. All errors are our own.

1 Introduction

We consider the problem of estimating demand for tied-good systems sold across competing retail formats. Goods and services are said to be tied when the purchase of a primary “tying” good requires consumers to purchase complementary “tied” or “aftermarket” goods from the same manufacturer. Tied-good systems are common in the economy: examples include printer cartridges tied to printers; after-sales services tied to durable-equipment purchases; and in the case of our application, razor-blades tied to specific razor technologies. The tying of two goods imposes a complementarity such that price increases for either good decreases demand for the other through its effect on the overall price for the system. Estimating the extent of the complementarity is key to several questions of interest to firms, including the optimal pricing of the primary and aftermarket products, compatibility decisions regarding primary and aftermarket goods, as well decisions regarding optimal bundle offerings. When tied-goods are sold in downstream retail channels, manufacturers also desire to understand the extent of complementarities between primary and aftermarket goods sold across competing retail formats, as well as the extent of substitution between competing systems of goods sold across those formats. Real-world market contexts typically feature competing manufacturers selling systems of tied-goods through oligopolistic, downstream retail formats. One of the goals of this paper is to present a parsimonious framework to measure the complex complementarity and substitution effects that arise in such contexts.

We define a forward-looking dynamic discrete choice model of demand that provides a realistic characterization of the two primary phenomenon of interest: tied goods and retail format substitution. Demand for most tied goods is inherently dynamic because the choice set and prices paid of aftermarket goods are contingent upon whether a primary good has been purchased in the past. A realistic perspective on how customers substitute purchases of a single good across retail outlets is also inherently dynamic if the good is durable and/or storable. Specifically, it is unlikely that customers know the prices of all products in a category at every store, then select the store with the lowest price. Rather, customers likely arrive at a store based on factors exogenous to the category such as distance or the price of focal products that drive store choice, then select to either stock up on the product at that store at the observed prices, or to postpone purchase in the category until a visit to a competing store with lower *expected* prices in the category. We therefore build the dynamics of our model around the choice set implications of tying and a stockpiling model that allows customers to endogenously determine the store at which they accumulate inventory.

Two features of the dynamic model are particularly relevant to both tying and retail substitution: usage rates and price expectations. We model usage of the aftermarket good to be endogenously determined based on consumer's current inventory levels, their price sensitivities, and the time until they expect to make another purchase to replenish their inventory. This aspect is critical for retail substitution because customers can slow their consumption to ease the process of waiting for a lower price at the current store or at a different store. Endogenous usage is also fundamental to the analysis of tied goods. Specifically, usage of aftermarket goods such as blades or printer cartridge is determined by how willing a customer is to use a depreciated blade or to read the fading ink from a nearly empty printer cartridge. The model of endogenous usage just described provides an intuitive way to relate this tolerance for depreciated goods to fundamental preference parameters such as a customer's price sensitivity. Less price sensitive customers should intuitively change their blades more often. In fact, many researchers have argued that because more sensitive customers buy fewer aftermarket goods, one way to price discriminate is to attract them through low primary good prices, while extracting more surplus from higher willingness to pay customers by charging a high aftermarket price. This form of price discrimination has been referred to as metering because the aftermarket good "meters" the intensity of demand for the system (c.f. Burstein 1960; Oi 1971; Schmalensee 1981; Klein 1996; Gil and Hartmann 2007).

Price expectations are a necessary component of any forward-looking dynamic demand analysis, but have unique implications for both retail substitution and tying. The recent empirical literature on stockpiling dynamics, c.f. Erdem, Imai and Keane (2003; henceforth EIK) and Hendel and Nevo (2006), has specified price expectations defined at the level of a single store or the entire market, in order to endogenize whether a customer should stockpile at current prices or wait for a future price discount. When retail substitution is considered, price expectations must be defined at each retail location a customer may possibly visit. In these contexts, the temporal variation of prices faced by the customer may be driven as much by the temporal variation in store visits as it is by temporal variation in prices at a given store. As described above, we model customers to only know the exact prices upon entering a store, but to substitute their purchases across stores based on the expected prices at competing stores and the likelihood of visiting a competing store.

Price expectations are also a well-recognized determinant of the pricing incentives for tied goods. Early antitrust inquiries into tying arose because regulators were concerned that high prices for tied aftermarket goods were the result of firms exploiting customers that were locked-in after the purchase of a primary good (e.g. Shapiro 1995; Hall 1997). The "Chicago School" critiqued this argument noting that when customers have rational expectations about future

aftermarket prices, a tie only allows a firm to shift extraction of a customer's willingness to pay from the primary good to the aftermarket good (for e.g. see, Carlton & Waldman, forthcoming). By including customer's price expectations in the model, we avoid overstating willingness to pay for the primary good and allow the price of the aftermarket good to affect the demand for the primary good. This yields substitution patterns consistent with the complementarity implied by the tie.

An important feature to consider when tied goods are sold through competing retail formats is the presence of alternatives to the tied good. Many classic tied good examples such as IBM computers and punch cards, and most of the theory literature, involves a monopolistic provider of the primary good. Yet retailers offering tied goods also often offer competing alternatives that may or may not be tied. For example, HP printers and cartridges are typically sold adjacent to Lexmark or Epson products, and Gillette Mach razors and blades share shelf space with other Gillette tied technologies, Schick razor and blade systems, and disposable razor blades. The willingness of customers to substitute to these other technologies is an important determinant of pricing at the level of both the manufacturer and the retailer.

To measure and analyze the substitution patterns implied by i) the complementary relationship between the tied goods, ii) the substitutability between a tied good and other alternatives, and iii) the substitutability between retail providers, we apply the model to individual-level panel data on consumer purchases of razors and blades at competing retail formats in a large Midwestern city in the US. The data span 56 weeks between 2002 and 2003. During this period, the market leader, Gillette, marketed the *Mach* as its flagship razor technology, while continuing to sell a substantial number of refill blades for its previous generation razor, the *Sensor*. The other significant competitor, Schick, had little market share compared to Gillette (the more successful Schick razor, the *Quattro*, had yet to be introduced). We observe all purchases of these two brands and disposable razor blades for every household across all store formats, including mass merchandisers and club stores. We focus on retail substitution at the format level because price expectations are likely to exhibit more variation between formats than between stores within a format. Our data also limit us to the format level because we are unable to match store level prices to the stores visited by consumers in our data.¹ An important caveat therefore is that any substitution we measure is likely to be conservative, providing lower bounds on retail pricing power, which is limited by additional inter-store substitution within retail formats.

¹ We use the publically available IRI store price files described in Bronnenberg et.al. (2008). However, our proprietary panel data comes from alternative sources that can only be matched to the IRI data at the format level. We chose not to use the Pittsfield and Eu-Clair panels in the Bronnenberg et.al. (2008) data because they do not include club stores, which sell substantially larger pack sizes at the lowest per-blade prices.

We estimate the model by maximum-likelihood. We accommodate a rich specification for unobserved heterogeneity via a flexible multivariate specification for the distribution of parameters in the population. The rich specification of random coefficients helps us sort the state dependence induced by inventory accumulation from persistence induced by permanent, unobserved heterogeneity (c.f. Heckman 1981). The control for heterogeneity complicates the estimation of the dynamic model, which is facilitated by the importance sampling method proposed by Akerberg (2001). Our results reveal evidence for strong complementarities in demand between razors and blades. The estimates suggest that manufacturers may profitably shift margins between the primary and aftermarket good, but cannot exploit customers in the aftermarket without suffering decreased primary good adoption. We also find evidence for strong retail substitution. Our estimates suggest that a permanent unilateral price increase at one retail format would lead to enough substitution to the other formats that a manufacturer would lose only a fraction of the sales lost by the retailer increasing prices. In spite of being conservative, the retail format competition we identify stands in contrast to much of the extant empirical literature that has either assumed monopolistic retailers, or found very limited retail competition.

We believe the latter conclusion in the extant literature has been driven by two aspects: first, a typical assumption is that customers know all prices at all stores. If customers do not know all prices, but rather respond to price expectations at competing stores as in our model, a lack of responsiveness to price realizations at stores not visited could be misinterpreted as a lack of cross-store substitutability. In effect, an assumption of perfect knowledge leads to a standard attenuation bias in estimation of cross-store price effects. Second, the extant literature has used static models that measure the *current*-period effects of prices changes. Cross-store price effects are likely to be long-term phenomena, wherein shifts in traffic and demand across stores occur over a period of time during which consumer expectations of price distributions across stores adjust in response to permanent price changes. These are best captured by dynamic models that formally incorporate a role for price expectations across stores.

The framework we present is most closely related to the empirical models of stockpiling dynamics proposed by EIK and Hendel and Nevo (2006). We extend their work in three ways. First, we allow for dynamics arising from the tied nature of the products. Second, we endogenize per-period consumption by explicitly embedding a nested, durable-good replacement policy into the solution algorithm. Third, we accommodate retail format choice in the stockpiling model, and endogenize the retail format at which consumers choose to accumulate inventory. The latter component is critical to measuring retail competition for durable or storable goods and constitutes what we believe, is a new approach for measuring retail substitution in single-category demand analyses. The frameworks we consider are also related to demand systems involving

complementary goods such as those in the indirect network effects literature, which considers hardware platforms bought in anticipation of future software purchases (e.g., Nair, Chintagunta & Dubé 2004; Dubé, Hitsch and Chintagunta 2007; Lee 2007).

Finally, in discussing the results of the paper, we identify two significant channel issues that manufacturers of tied goods face when selling through an oligopolistic downstream channel. We note that double marginalization may be particularly problematic when selling tied goods due to the common desire of manufacturers to maintain low primary good prices. Furthermore, we identify what we call a “cross-product” horizontal externality that arises because the pricing incentives of the retailer, who cannot tie sales of aftermarket products to his store, are different from those of the manufacturer. If there is limited retail competition, the availability of aftermarket products at substitute stores may provide an incentive to a retailer to set primary good prices that are “too high” from the manufacturers perspective. The potential for these channel problems are relevant for managers and depend on the level of retail substitution, which our framework can be used to measure. In this respect, we believe our analysis motivates theoretical inquiry into the previously unstudied problem of selling tied goods through an oligopolistic retail channel.

The rest of the paper proceeds as follows. In the next section, we present our empirical model of demand. The next section presents the data on consumer purchases of razors and blades used in the study. We then present parameter estimates from the model, and then discuss the implications for retailers and manufacturers. The last section concludes.

2 Model

This section develops a dynamic model of demand for tied goods sold through competing distribution channels. The formulation of the empirical model is motivated by the considerations discussed in the introduction, as well as some of the specifics of our application to razors and blades.

Market

There are $R+1$ shaving options in the market, available across K different retail outlets. The options comprise of R tied razor technologies, and one non-tied option, i.e., disposables. The tied razor technologies require the purchase of a razor before the corresponding blades can be used. Blades of a given razor technology are incompatible (by design) with razors of other technologies. We index the tied razor technologies by $r = 1, \dots, R$, and let $r = 0$ index disposables. We also incorporate the institutional feature that razors are always sold in packs containing blades. For

each tied razor technology, we let $j = 1$ index the pack that contains the razor, and index the remaining packs that contain only blades by $j > 1$. We denote the total number of pack sizes available for razor technology r at retail outlet k , as J_{rk} . Disposables are available in J_{0k} different packs, where each pack $j \in \{1, \dots, J_{0k}\}$ only contains disposable razors. Each pack, whether containing a razor or only blades, contains q_{rjk} blades.

In our empirical context $R = 3$, corresponding to Gillette-Mach, Gillette-Sensor, and Schick brands of razors. Our empirical model considers three retail types, corresponding to *DG* (Drug/Grocery), *MK* (Mass Merchandisers and Super Centers) and *W* (Wholesalers and Club) stores. Thus, $K = 3$.²

Consumers and States

At the beginning of each week, a consumer owns a set of razors, a set of unused blades compatible with one of these razors, and a used blade that has depreciated from its initial quality-level due to usage. In addition to these, the consumer's state also incorporates the retail format visited and the prices at that outlet. We describe each of these in sequence below.

Razor technology ownership (Ω):

A consumer's razor technology ownership status at the beginning of each week is denoted by Ω , which is an $R \times 1$ vector of indicators for whether or not each razor type is owned. Once a razor, $r = 1, \dots, R$, is purchased, the consumer is assumed to have it forever, and element r of Ω is permanently set to 1.³ An implication of this assumption is that consumers can own multiple razor technologies at the same time (i.e. multi-homing). The data suggests multi-homing since we observe several instances of consumers switching between blade packs of different brands without purchasing razors in between (see Appendix A).

Blade type currently owned (ρ):

² We limit our analysis to three distinct retail formats to manage the computational tractability of the model. We treat club stores as distinct due to their unique pack sizes, but merge mass merchandisers and super centers given their similar price levels. Though drug and grocery stores differ in their layout and perhaps the type of visit, we choose to merge them since the prices and pack sizes offered in these stores are very similar in this category.

³ The perfect durability assumption is motivated by the fact that we rarely observe the same household buying the same razor brand more than once during the time-period of our data. Further, we also observe patterns in the data where consumers are observed buying blades of a given brand on purchase occasions that are far apart in time, without purchases of any associated razors of that brand in between. This suggests that our no disposal assumption regarding razors ownership is reasonable. It is possible to extend the model to allow for depreciation in the quality of the primary good in other contexts where that may be more relevant. An alternative model is to assume that consumers dispose the primary good (i.e., razors) every time they switch technologies (i.e., $\rho' \neq \rho$). This assumption makes the model simpler, but has to be evaluated on a case by case basis, depending on the category being studied.

At the beginning of a week, a consumer also owns blades of a type $\rho \in \{0,1,\dots,R\}$. Reflecting the nature of the market, a consumer can own disposable razors irrespective of his razor technology ownership status (i.e., ρ can be 0 irrespective of Ω), but can own non-disposable razor blades only if a corresponding razor is available (i.e., $\rho = r > 0$ only if $\Omega_r = 1$). The table below provides three examples to clarify the notation (e.g., in row three below, the consumer owns all three razors and is currently using disposable razors).

Ω	ρ	Razor technologies owned			Blade type currently owned			
		Mach	Sensor	Schick	Disposables	Mach	Sensor	Schick
(0,1,1)	0		×	×	×			
(1,1,1)	1	×	×	×		×		
(1,1,1)	0	×	×	×	×			

Thus, a consumer can own many razor technologies at the same time, but can own only one type of blades. We make the latter assumption to keep the empirical model tractable. Relaxing this assumption would require us to add the stock of blades of each type as a state variable, and to model the endogenous decision of which blade stock to consume from.⁴

Level of depreciation (z):

We allow blades to depreciate in quality with usage (we explain how we model the depreciation rate in more detail below). The utility from usage of a blade is highest when it is new and lowest when it is dull. We capture the level of depreciation of a blade by modeling the fraction of initial per-usage utility that the blade can currently provide. Letting $z \in [0,1]$ denote the depreciated level of the blade currently used, we model the per-usage utility from the current blade as γz , where γ denotes the utility from usage of the blade when new.

Stock of unused blades (b):

We let b denote the stock of *unused* blades available in inventory with the consumer (i.e., total blades in inventory, including the current depreciated blade, is $b+1$). If the consumer has no unused blades in inventory, $b = 0$, and the current blade is completely dull, $z = 0$, the consumer may be using the completely dull blade or some other technology (e.g., electric razor) for shaving. This is the outside option for the consumer.

⁴ In the related dynamic stockpiling literature, past researchers have also avoided modeling multiple inventory types and the associated endogenous decision of which type to consume by assuming there is a single stock. Hendel and Nevo (2006) assume all brands are identical post-purchase, while EIK allow the quality of a single stock to be indexed by the proportion of each product making up the stock. While these assumptions fit the categories they consider, both seem inappropriate in the razor blade tied good context where we expect agents to prefer to consume the type of blade most recently purchased. Our assumption that agents can own only blades of the most recently purchased razor better captures this aspect of consumer behavior.

Store visited (k):

Each week, the consumer visits a retail outlet from one of the K different channels, $k = \{1, \dots, K\}$, or stays at home. The decision to not visit any outlet is denoted $k = 0$.

Prices (\mathbf{p}_k):

At the visited outlet, k , the consumer observes a price, p_{rjk} , for each of the razor and blade packs available. The vector of prices at outlet k is \mathbf{p}_k , where $\mathbf{p}_k = \{p_{11k}, \dots, p_{R, Jrk}\}$.

Together, these state variables form the state vector $\mathbf{s} = \{\Omega, \rho, z, b, k, \mathbf{p}_k\}$.

Actions

There are two types of actions in the model. The first is a decision involving *purchase* of razors or blades at a visited store, and the second is a decision involving *replacement* of the currently used blade at home with a new one from the available inventory. Given state \mathbf{s} at the beginning of the week, the consumer makes one of the following three purchase decisions.

- a. Switch blade type to $r \neq \rho$ by buying pack $j^* \in \{1, \dots, J_{kr}\}$
- b. Stay with the current blade type ρ and not buy any blades, or,
- c. Stay with the current blade type ρ and buy the blade pack $j^* \in \{1, \dots, J_{k\rho}\}$ that gives him the best payoff

If the consumer decides to stay with the current technology, the consumer also decides whether or not to replace his current blade or continue using it at its depreciated level. The blade-replacement decision determines the consumer's consumption rate and is endogenous in our model. We model the blade-replacement decision by incorporating a nested durable good replacement problem of the sort considered by Rust (1987) or Gordon (2007) for the storable goods (see Sun 2005 for an approach which treats consumption as continuous). This decision occurs at the end of a period and affects the level of depreciation of the blade used in the next period. The blade replacement decision is therefore a component of the state transitions presented below.

Each week, the consumer chooses the actions that maximize the expected present discounted value of current and future payoffs associated with those actions.

Utilities

We define utility over the usage of blades. The utility from *purchase* of a razor or blades is the discounted present value of the expected stream of utility from subsequent consumption of blades. Below, we present the immediate utilities associated with each of the three purchase decisions, then later define the expected future utilities that also affect purchase incentives. We define

parameters and utilities for a single agent, but allow them to be heterogeneous across agents in our estimation.

Current period utility of switching blade type to $r \neq \rho$

The consumer can switch his current blade type from ρ to r by buying a pack of type r from any of the K retail formats in the market. If a consumer switches his blade type in a given week, we assume he consumes a new blade that week, and thus receives flow utility from consumption, γ_r , (since $z = 1$). We then write the current utility from *purchase* of pack j of type $r \neq \rho$ as,

$$u_{rj}(\mathbf{s}) = \gamma_r - \alpha p_{rjk} + \kappa + \lambda \gamma_\rho b + \varepsilon_{rj} \quad (1)$$

Here, α is the price sensitivity, p_{rjk} is the price of the pack, and ε is an iid extreme-value shock to utility, unobservable to the econometrician. The term κ is a (dis)utility from making any purchase, which could be interpreted as a carrying cost or transaction cost realized at the point of purchase. Econometrically, κ helps fit the infrequent incidence of purchase of razors and blades across observed store trips. We include the term $\lambda \gamma_\rho b$ to allow the consumer to receive a salvage value for the stock of blades from the old technology, ρ . We model the salvage value as a fraction of the utility that the consumer may have attained from consuming the b blades he had available in inventory, when deciding to switch. λ can therefore be interpreted as the share of the consumer's utility for a blade that is received when the consumer sells it in a resale market or the discounted value of using that blade at some distant future date. The salvage value serves three aspects in the model. First, the economic switching costs in this category should be bounded by the cost of a razor, which is small. A "free disposal" assumption under which a user switching razors disposes all existing blades without realizing any value, artificially increases switching costs. By allowing for a salvage value from old blades when replacing razor types, we partly address this concern. The salvage value term also generates a computational advantage since it allows us to keep track of a single stock of blades. Finally, the salvage value term can improve the fit of the model since it allows the stock of past blades to directly enter the per-period utility function for switching blades.

Current utility of staying with current blade type ρ and not buying any blades

If a consumer stays with the current blade type ρ and chooses to not buy any blades, he obtains the per-period utility of using his current blade:

$$u_0(\mathbf{s}) = \gamma_\rho z + \varepsilon_0 \quad (2)$$

If the current blade is completely depreciated, $z = 0$, and $b = 0$, the outside option is consumed and its utility is normalized to zero.

Current utility of staying with current blade type ρ and buying blade pack j

Finally, the consumer can decide to stay with the current blade type ρ and buy a pack of compatible blades to refill his inventory. In this case, the consumer obtains current utility:

$$u_{\rho j}(\mathbf{s}) = \gamma_{\rho} z - \alpha p_{\rho j k} - \kappa + \Delta + \varepsilon_{\rho j}, j > 1 \quad (3)$$

In contrast to switching blade types, the consumer does not receive a salvage value when buying refills. We allow for a term Δ , which represents an increase in utility from buying the same type of blades as in the past. This is a form of habit persistence likely the result of a psychological cost of switching brands (e.g. Roy, Chintagunta and Halder 1996; Dubé, Hitsch and Rossi, 2008).

State Transitions

Dynamics arise in this model through the transitions of the state variables $\{\Omega, \rho, z, b, k, \mathbf{p}_k\}$. The key dynamics of interest are generated by the state variables Ω, ρ, z, b . A choice in the current period affects the razor technology and stock of blades possessed next period, as well as the level of depreciation of the blade to be used next period. To economize on notation in what follows, we denote the next period's state by an apostrophe.

Denote the set of possible purchase decisions at retail outlet $k > 0$ as $\mathbf{y}_k = \{y_{k0}, (y_{k01}, \dots, y_{k0j_{0k}}), \dots, (y_{kR1}, \dots, y_{kRj_{Rk}})\}$. y_{k0} is an indicator variable denoting no purchase (of razors or blades) while at store k ; the indicators y_{krj} ($r \neq \rho$) denote purchase of pack j for a new technology r ; and the indicators $y_{k\rho j}$ denote purchase of blade pack j of the currently owned blade type ρ . Finally, y_0 is an indicator of no store visit. Due to the discrete choice aspect of the problem, we have that,

$$y_0 + \sum_{k=1}^K y_{k0} + \sum_{k=1}^K \sum_{r=0}^R \sum_{j=1}^{j_{rk}} y_{krj} = 1$$

We now present the transitions of the state variables.

Transition of razor ownership (Ω)

The law of motion for razor ownership status is governed by the purchase decision:

$$\Omega'_r = \begin{cases} 1 & \text{if } y_{kr1} = 1, r > 0, k > 0 \\ \Omega_r & \text{otherwise} \end{cases} \quad (5)$$

Thus, purchase of any pack $j = 1$ of a non-disposable razor technology adds r to the set of razors owned. The above specification implies that $\Omega_r = 1$ is an absorbing state.

Transition of current blade type (ρ)

The law of motion for the blade type currently owned is governed by the purchase decision:

$$\rho' = \begin{cases} r & \text{if } \sum_j y_{krj} > 0, r \neq \rho, k > 0 \\ \rho & \text{otherwise} \end{cases} \quad (6)$$

Thus, purchase of any pack, whether containing a razor plus blades, or only blades of a new technology r , switches ρ to r .

Transition of level of depreciation (z)

At the beginning of each period, the consumer's current blade is at level z . If the consumer decides to stay with the current razor technology, but replace the current blade, the level next period will be equal to its maximum value, 1. If the consumer decides to stay with the current razor technology, but not replace the current blade, we assume that the level will be depreciated to δz , where δ is a parameter to be estimated from the data. Finally, if a consumer changes his blade type in that period, a new blade will be used immediately, such that that next period's blade is of level δ . The transition equation is therefore:

$$z' = \begin{cases} 1 & \text{if } \rho' = \rho, c = 1 \\ \delta z & \text{if } \rho' = \rho, c = 0 \\ \delta & \text{if } \sum_j y_{krj} > 0, r \neq \rho, k > 0 \end{cases} \quad (7)$$

In (7), c is an *endogenous indicator* for whether or not the consumer changes blades. The choice of c will be described once we define the value functions below. In the empirical application below, we allow the depreciation rate δ to be different for disposables and non-disposables.⁵

Transition of stock of unused blades (b)

The transition function for the blade stock depends on whether a consumer switches his blade type, purchases refills for an existing razor, or does not make a purchase. If the consumer switches his blade type, the consumer's stock of *unused* blades next period equals the number of blades included in the pack minus 1. We subtract one because a consumer immediately starts

⁵ An alternative assumption to capture the replacement decisions would be to directly estimate the weekly consumption of blades as a parameter, and to depreciate the stock of blades in inventory by this parameter. Relative to this approach, our specification has the advantage of endogenizing weekly consumption, and also of capturing the discrete nature of replacement decisions in the blade category. The latter aspect likely better approximates the actual process of replacement decisions in this, and many other tied good categories. For example, in the case of printers and cartridges, an archetypal tied good market, z captures the level of depreciation of the cartridge (how much ink is left); in the market for Polaroid cameras and films, z captures the level of depreciation of the film currently used (how many photos are left?) Once the ink runs out, or all the photos are taken, a new cartridge or roll of film is used, and the stock is depleted by one, as in our model.

using one of the new blades when purchased. Old blades, b , as described earlier are not retained. When a consumer stays with the current blade type, $\rho' = \rho$, the stock is depreciated by 1 if he chooses to replace the current blade (i.e., if $c = 1$), and is augmented by the number of blades corresponding to the purchase decision, $q_{\rho jk}$:

$$b'(\rho, y|c) = \begin{cases} q_{rjk} - 1, & \text{if } y_{krj} = 1, r \neq \rho, k > 0 \\ b - c + q_{\rho jk} & \text{if } y_{\rho jk} = 1, \rho' = \rho, k > 0, (b - c + q_{\rho j}) \leq \underline{B} \\ \underline{B} & \text{if } y_{\rho jk} = 1, \rho' = \rho, k > 0, (b - c + q_{\rho j}) > \underline{B} \end{cases} \quad (9.1)$$

Note that a consumer can only change blades (i.e. $c = 1$) if the current stock of unused blades is greater than zero ($b > 0$) and/or he purchases blades ($q_{\rho jk} > 0$). We restrict the number of blades to be less than \underline{B} to keep the state space for the dynamic problem bounded. This is not a binding constraint because we set \underline{B} such that it is greater than or equal to the maximum inventory a consumer would hold.

Store visited

We assume that each consumer has a probability ϕ_k of visiting outlet k each week. We account for persistence in retail outlet visits by allowing the consumers' past retail outlet visit decisions to shift ϕ_k . Given that razors and blades form a small part of the overall shopping budget, our implicit assumption is that the store visit probabilities are not shifted by prices or consumer inventories (we find support for this assumption in the data). The current model would have to be extended to accommodate product categories which may actually drive store choice decisions (e.g. Bell and Lattin 1998). However, the decision of which retail format to purchase and accumulate inventory at is endogenous in the model, and is the outcome of the dynamic choice. Letting $I(k)$ denote an indicator for whether the consumer visits retail outlet k , we write the outlet visit probabilities as,

$$\phi'_k = \Pr(I'(k) = 1) = \frac{\exp(\mu_k + \nu I(k))}{\sum_{m=1}^K \exp(\mu_m + \nu I(m))} \quad (10)$$

where, ν and μ_k are parameters to be estimated from the data.

Prices (p)

A flexible specification of the transition density for prices would allow for prices that are correlated across products in a given week, and are serially-dependant over time. With a large number of products, it is infeasible to estimate parametric specifications with unrestricted variance-covariance matrices that capture the co-movement in product prices flexibly. Instead, we

allow each price to have two support points (i.e., a base price and discount price, as frequently observed in retail scanner panel data). There are consequently $2^{\sum_r J_{rk}}$ possible combinations of prices at each outlet. We assume that consumers expect each of these price-combinations to be realized with the probability with which we observe it in the data, ψ_{km} where m denotes a given combination of prices (in our empirical application, m is of the order of 512 at the drug/grocery channel). We believe this provides a reasonable approximation to how consumers form expectations over prices, but it should be noted that the actual price realizations we analyze in the data will not be strictly binary.

Implicitly, store format prices are allowed to be flexibly contemporaneously correlated across products, but assumed to be independent over time. A correlation in realized prices *seen* by the consumer does however arise through the state dependence in the store visit probabilities. For instance, a consumer may anticipate during a club store visit today that his chance of visiting a club store again in the immediate future is small. Hence, expected prices seen tomorrow are higher, since tomorrow's visits are more likely to other, higher-priced formats.

Value Functions

The laws of motion of the state variables imply that consumer's current actions affect their states in the future. Hence optimal decision-making involves picking the alternative that maximizes the consumer's "alternative-specific" value function (Rust 1987). In the remainder of this section, we derive the alternative-specific value functions separately for whether or not the consumer switches to a new blade type.

Value of switching to blade type $r \neq \rho$ by purchase of pack j

We write the value function for switching to a new blade type as,

$$\begin{aligned}
 V_{rj}(\rho, z, b, \Omega, \mathbf{p}_k) = & \underbrace{(\gamma_r - \alpha p_{rjk} + \lambda_\rho \gamma_\rho b)}_{\text{I}} + \underbrace{\beta \phi_0 \mathbf{E} \left[V_0(r, \delta, b'(r, y), \Omega') + \varepsilon_0 | b, \Omega \right]}_{\text{II}} \\
 & + \beta \sum_{k>0} \phi_k \sum_m \psi_{mk} \mathbf{E} \max \left\{ \underbrace{V_0(r, \delta, b'(r, y), \Omega') + \varepsilon_0}_{\text{III}}, \underbrace{\left\{ V_{rj}(r, \delta, b'(r, y), \Omega', \mathbf{p}_k) + \varepsilon_{rj} \right\}_{j=1}^{J_{rk}}}_{\text{IV}}, \right. \\
 & \left. \underbrace{\left\{ V_{\tilde{r}j}(r, \delta, b'(r, y), \Omega', \mathbf{p}_k) + \varepsilon_{\tilde{r}j} \right\}_{j=1}^{J_{rk}, \tilde{r} \neq r}}_{\text{V}} \right\}
 \end{aligned} \tag{11}$$

The first term (I) in the value function above captures the current period utility of switching to razor technology r today via a purchase at retail outlet k . The second term (II), captures the option value of potentially making no outlet visit (with probability ϕ_0) and staying on with the

new blades r tomorrow. We define V_0 below. A switch to a new blade type today also encapsulates the discounted value of potentially visiting each outlet k with probability ϕ_k tomorrow, and choosing to either make no purchase (term **III**); buying refill blade pack j for the new technology (term **IV**); or switching to a new blade type $\tilde{r} \neq r$, via purchase of a pack $j \in (1.., J_{\tilde{r}k})$ (term **V**). The option value of an outlet visit is the expected maximum of these actions, integrated over the set of future prices possible at the outlet, \mathbf{p}_k , and over the distribution of unobservable shocks to utility that could be realized in the future (i.e., ε -s), evaluated at the value of the other state variables tomorrow $\mathbf{s}' = \{\Omega', \rho' = r, z' = \delta, b' = b'(r, j), k, \mathbf{p}_k\}$, and discounted by one period by the factor β .

Value of staying with current blade type ρ

The value to a consumer from staying with the current blade type depends on the choice of c . To choose c , the consumer compares the discounted future value of changing the blade with that from continuing to use the current blade. We therefore begin by specifying the discounted future value under both scenarios. If he changes the blade currently used, he has the following expected future payoff,

$$EV_{\rho_j}(\rho, z, b, \Omega | c = 1) = \beta \phi_0 \mathbf{E} \left[V_0(\rho, 1, b'(\rho, y | c = 1), \Omega) + \varepsilon_0 | b > 0 \right] \\ + \beta \sum_{k > 0} \phi_k \sum_m \psi_{mk} \mathbf{E} \max \left\{ \begin{array}{l} V_0(\rho, 1, b'(\rho, y | c = 1), \Omega) + \varepsilon_0, \\ \left\{ V_{\rho_j}(\rho, 1, b'(\rho, y | c = 1), \Omega, \mathbf{p}_k) + \varepsilon_{\rho_j} \right\}_{j=1}^{J_{\rho k}}, \\ \left\{ V_{\tilde{r}_j}(\rho, 1, b'(\rho, y | c = 1), \Omega, \mathbf{p}_k) + \varepsilon_{\tilde{r}_j} \right\}_{j=1}^{J_{\tilde{r}k}, \tilde{r} \neq \rho} \end{array} \right\} | b > 0 \quad (12)$$

If he stays with the current blade, the consumer starts off the next period with $z' = \delta z$, with expected future payoff,

$$EV_{\rho_j}(\rho, z, b, \Omega | c = 0) = \beta \phi_0 \mathbf{E} \left[V_0(\rho, \delta z, b'(\rho, y | c = 0), \Omega) + \varepsilon_0 \right] \\ + \beta \sum_{k > 0} \phi_k \sum_m \psi_{mk} \mathbf{E} \max \left\{ \begin{array}{l} V_0(\rho, \delta z, b'(\rho, y | c = 0), \Omega) + \varepsilon_0, \\ \left\{ V_{\rho_j}(\rho, \delta z, b'(\rho, y | c = 0), \Omega, \mathbf{p}_k) + \varepsilon_{\rho_j} \right\}_{j=1}^{J_{\rho k}}, \\ \left\{ V_{\tilde{r}_j}(\rho, \delta z, b'(\rho, y | c = 0), \Omega, \mathbf{p}_k) + \varepsilon_{\tilde{r}_j} \right\}_{j=1}^{J_{\tilde{r}k}, \tilde{r} \neq \rho} \end{array} \right\} \quad (13)$$

Equations 12 and 13, jointly define a *consumption policy function* that determines the optimal consumption of blades as a function of the consumer's state. The optimal consumption policy is nested into the computation of the choice-specific value functions. The expected value of the

consumer's best possible action is the maximum of the two payoffs in (12) and (13). We can therefore write the value function for staying with the current razor as:

$$V_{\rho_j}(\rho, z, b, \Omega, \mathbf{p}_k) = \begin{cases} (\gamma_{\rho} z - \alpha p_{\rho_{jk}}) + \max_c \begin{cases} EV_{\rho_j}(\rho, z, b, \Omega | c = 1), \\ EV_{\rho_j}(\rho, z, b, \Omega | c = 0) \end{cases} & \text{if } b > 0 \text{ or } q > 0 \\ (\gamma_{\rho} z - \alpha p_{\rho_{jk}}) + EV_{\rho_j}(\rho, z, b, \Omega | c = 0) & \text{if } b = 0 \text{ and } q = 0 \end{cases} \quad (14)$$

Note that if the customer has no unused blades, $b = 0$, and does not make a purchase (y_{k0} or y_0 equals 1), the customer cannot change blades (i.e. $c = 0$). The value when not visiting a store, V_0 , is defined analogously to (14), except the current period utility excludes a price effect because nothing is purchased. In essence, (14) incorporates a durable-good replacement problem analogous to Rust (1987), for the blades owned. An appealing feature of this formulation is that replacement is endogenous, driven by the consumer's preferences, as well his expectations about future prices and inventory. This captures the commonly noted phenomenon that individuals may use more blades when there are many in inventory, yet retain the current blade longer when the inventory is reduced. Figure 1 illustrates how the endogenous blade replacement is affected by current inventory and the depreciated value of the current blade, for one of the consumers in the data. We see that when the consumer has 60 blades in inventory, the value of replacing (equation 12 above) is always greater than the value of retaining the current blade (equation 13). However, at inventory levels of 20, 8 and 4, the customer increases his use of the existing blade to depreciation levels of 0.75, 0.37, and 0.29 before replacement.

--- Figure 1 here --

Maximum likelihood estimation

We estimate the parameters of the model via maximum likelihood. We now add the subscript i denoting consumer, and t denoting "week" to all variables. We collect the parameters specific to a consumer in a vector, $\theta_i = \left\{ (\gamma_{ir})_{r=0}^R, \delta_{i,disp}, \delta_{i,nondisp}, \alpha_i, \lambda_i, (\phi_{ik})_{k=1}^K, \nu_i \right\}$ (here, we allow the depreciation rate δ to vary between disposables and non-disposables). The discount factor, β , is not estimated, and is set to 0.998.⁶

Heterogeneity

⁶ This is roughly equivalent to a 10% annual interest rate.

We allow for a multidimensional continuous heterogeneity distribution using random coefficients. We specify the parameters, θ_i , to be multivariate normally distributed across consumers as $\theta_i = \bar{\theta} + \Gamma \eta_i$, where Γ is the Cholesky decomposition of the covariance matrix of the parameters, Σ , and η is a vector in which each element is distributed standard normal (while all parameters are modeled to have a normal underlying distribution, some parameters are transformed to either restrict them to have positive or negative support, or to bound them between 0 and 1, as is necessary for δ). We denote the set of parameters to be estimated as, $\Theta = \{\bar{\theta}, \Gamma\}$. Given the extreme-value assumption on the stochastic terms to utility, the probability of purchase for each week t is given by a logit function in the alternative-specific values (please see Appendix D for exact formulae). The likelihood function for a given individual in the data for a week t is therefore:

$$\begin{aligned} \ell_t \left(y_{i0t}, \mathbf{y}_{i1t}, \dots, \mathbf{y}_{iKt} \mid \rho_{it}, z_{it}, b_{it}, k_{it}, \Omega_{it}, \mathbf{p}_{kt}; \theta_i \right) \\ = \left[\Pr(y_{i0t} = 1) \right]^{y_{i0t}} \prod_{k=1}^K \left[\Pr(y_{i0kt} = 1) \right]^{y_{i0kt}} \prod_{j=1}^{J_{\rho_t k}} \left[\Pr(y_{i\rho_t jkt} = 1) \right]^{y_{i\rho_t jkt}} \prod_{r \neq \rho_t} \left[\prod_{j=1}^{J_{rk}} \Pr(y_{irjkt} = 1) \right]^{y_{irjkt}} \end{aligned} \quad (18)$$

The likelihood of the data for the individual across all weeks is then,

$$L_i(\Theta) = \int \left[\prod_{t=1}^T \ell_t \left(y_{i0t}, \mathbf{y}_{i1t}, \dots, \mathbf{y}_{iKt} \mid \rho_t, z_t, b_t, k_t, \Omega_t, \mathbf{p}_{kt}; \theta_i \right) \pi(\rho_{i0}, z_{i0}, b_{i0}, \Omega_{i0}) \right] d\Phi(\theta_i; \Theta) \quad (19)$$

where ρ_{i0} , z_{i0} , b_{i0} , Ω_{i0} are the initial current razor, current blade depreciation level, blade stock, and razor ownership, and $\pi(\rho_{i0}, z_{i0}, b_{i0}, \Omega_{i0})$ is the probability of the initial condition. An obvious issue to consider in constructing this likelihood is that b_t and z_t are not observed in the data. Consequently, for each guess of the parameter vector, we infer these according to the state transitions defined above, the model parameters, the solution of the dynamic programming problem, the blade purchases in a given week, q_{ijt} , and the initial states.⁷ The initial states are unobserved, which creates a standard initial conditions problem. We derive the initial states based on model parameters and the weeks until the consumer is observed to make his first purchase.

⁷ As in other storable good models with unobserved state variables, we construct the likelihood assuming that we know the unobserved state with certainty. For example, if a consumer randomly decides not to shave for a few days, the current blade will not actually depreciate even though we assume it does. Ideally, one would integrate over all possible values of the unobserved states to account for random variation in depreciation, but this is significantly burdensome. Using our best prediction of the unobserved state as specified by the model we are able to model interpurchase times in a functional form of past purchases that, while not exact, is consistent with the behavioral primitives of interpurchase timing.

Specifically, to obtain the initial razor ownership Ω_{i0} , we assume that all consumers own Schick and Sensor razors from the beginning of the data. This assumption is motivated by the fact that while we see very few observations of purchases of razors of these technologies in our data, we nevertheless observe many purchases of blade packs of these technologies. We also add Mach to initial razors owned if the first Mach purchase observed for that consumer is a blade pack that does not contain Mach razors. We obtain the initial blade type ρ_{i0} as follows. If the first observed purchase for a consumer is a blade pack, we set ρ_{i0} to the type of that blade pack. If the first observed purchase for a consumer is a razor-plus-blade pack, we assume that the initial blade type owned is a disposable. We recompute the initial stock b_0 for each consumer for each guess of the parameters, based on a customer having an approximate stock of zero in the first week we observe him purchasing any packs. Specifically, if a consumer i is observed to make his first purchase T_i weeks after the beginning of the data, we assume that the initial stock is $T_i(1-\delta)$, where δ is the current guess of the depreciation rate for the consumer.⁸ Finally, we assume consumers start with a new blade at the beginning of the data – i.e., we set $z_{i0} = 1$.

The rich specification of heterogeneity increases the computational burden of the estimator significantly. To estimate random coefficients in this dynamic model we employ the change of variables and importance sampling technique proposed by Akerberg (2001). In a first step, we solve the value functions over a set of 2000 simulated parameter values. In subsequent iterations, we use this initial solution to obtain the likelihood at the new guesses of the parameter vector while avoiding the computational burden of resolving for the value functions.

Identification

We now present a brief informal discussion of identification in this model. The availability of panel data, including both the product choice as well as the quantity choices (i.e., pack sizes) of consumers, facilitates identification of preferences and heterogeneity. The usage utility parameters of each consumer, γ_i , are identified from his average propensity to purchase in the data, in addition to the total number of units of a given razor type that are observed to be purchased across all time periods. The incidence of temporary price cuts in scanner data provides price variation to identify price sensitivities. The salvage value fractions, λ , are identified from the time-elapsd as well as the pack sizes that are bought prior to purchase occasions that involved a change of blade types (i.e., holding consumption and usage utilities fixed, if we see a switch to a new blade type occurring soon after a large pack of a different blade type was bought,

⁸ Recall that if $\delta = 0$, a new blade will have to be used the next period. If $\delta = 1$, no replacement is necessary since the current blade has no depreciation. Hence, the expected blades used per period is roughly $(1-\delta)$.

we infer that λ is high). Consumption is unobserved in the data, and is identified by the joint distribution of interpurchase times and quantity choices, i.e., large pack size choices interspersed with short interpurchase times imply high consumption rates. We also allow for heterogeneous depreciation rates of stockpiles on multiple razor types. These are intuitively identified in the following way. If we see that the time-to-purchase of disposable blade packs following a disposables purchase is small, we infer depreciation of disposables is high. If we see that time-to-purchase of blade packs following a non-disposable blade pack purchase is small, we infer depreciation of non-disposable razor blades is high. Finally, the frequency of observed outlet visits for a given consumer identifies his outlet visit probabilities. Persistence in outlet visit probabilities for the consumer is identified from any persistence in the temporal distribution of his observed outlet visits.

3 Data & Descriptive Analysis

Our data contains the complete purchase history of razors and blades for a panel of consumers in a large Midwestern city in the US. An advantage of the data is that all purchases of products in this category at every retail outlet in the city are observed. To avoid confounds related to purchases of razors and blades for different members of the household, we restricted attention to purchases of only male razors and blades. Our final dataset contains the purchase patterns of male razors and blades for a panel of 725 consumers over a period of 56 weeks beginning in mid-September 2002 and ending in the beginning of October 2003. These data were augmented with a price file from the IRI dataset (c.f. Bronnenberg et. al. 2008) that provided prices for all retail formats in the city (except Club stores).

The period of our data includes the Gillette Mach razor, the Gillette Sensor razor, as well as the Schick Xtreme3 and Tracer (the data does not include the Schick Quattro which was introduced in late 2003). In the data, we observe several purchases of Mach razors and blades, but observe little or few purchase of Gillette Sensor or Schick razors, in spite of observing blades bought. Hence, while we model blade demand for all brands, we model *razor* adoption of only Mach; Sensor and Schick razors are assumed to be owned prior to the beginning of the data. The reader is pointed to Appendix B which discusses these issues extensively.

Data cleaning

In order to obtain a dataset suitable for estimation, we undertook several steps to clean the raw data. A detailed discussion of the data-cleaning steps is provided in Appendix B, available online from the journal website. We briefly summarize the steps below. First, we restrict attention to

households that purchased at least one male razor or blade pack during the period of the data. Second, for parsimony, we combine some of the infrequently purchased packs with commonly purchased pack-sizes. Third, we also model consumers to make a discrete choice over the channel visited. We use specific thumb-rules to assign consumers observed to be visiting multiple channels in the same week, to a single format. Finally, we do not have actual price charged at club stores, since these stores do not share data with syndicated vendors. Thinness of purchase data precluded inferring these weekly from observed purchases. Instead, we compute the mean price paid of each product across all consumers and across all weeks in the data, and use these as a proxy for the prices not observed. This methodology is admittedly imperfect, and is a limitation of the paper.

Summary Statistics of Individual-Level Data

We first describe the distribution in the raw data of the implied state variables from the model. Recall that in our framework, we model a customer beginning each week with a current razor technology and visiting one of the 3 store types or staying home. We summarize the incidence of each of these in Table 1.

--- **Tables 1 and Figure 2 here** ---

The bottom panel of Table 1 presents the distribution of razor ownership across consumers and weeks. Interestingly, about 31% of consumer-weeks involve ownership of a Gillette Mach razor, while only 29% of consumer-weeks involve ownership of Mach blades in inventory, suggesting that there are some instances in which a consumer chooses to use blades of another technology while owning a Mach razor. This partly motivates our model formulation that allows multiple razor ownership by consumers. Figure 2 plots the distribution of the total blades purchased in all months in the data across customers. This is an indirect measure of the distribution of consumption rates, and suggests significant heterogeneity.

Stocking and Sales of Pack Sizes by Channel

In Table 2, we list the pack sizes available at each retail outlet, as well as the prices of each of the packs by channel. Our data includes 15 pack types across 3 retail outlets. The grocery and drug and mass merchandiser channels carry similar packages, while the club stores only sell Gillette brands or disposables and only carry large package sizes. The differences in product availability

across differing retail channels provides an incentive for consumers of differing tastes to self-select into visiting and accumulating inventory at each retail channel.⁹

--- **Table 2 here**---

In general, Mach blades tend to be priced higher. Consistent with the notion that primary goods (i.e., razors) are priced cheaply, we do not see much evidence for higher prices on packs that contain razors. Table 2 also provides information in the differences in prices across retail formats. Consistent with the conventional wisdom, comparable pack sizes are priced lower at Mass merchandisers than at grocery stores. The club stores stock only large pack sizes which have low prices on a per-blade basis, which make them more attractive to high-consumption consumers. We also see some broad evidence for pack size based price discrimination via quantity discounts in this category. This aspect of price discrimination is accommodated in our empirical model since we explicitly model choices over pack sizes.

Descriptive Statistics about Channel Switching

An important piece of our analysis involves the measurement of retail substitution. Our model specifies retail substitution to occur through consumers endogenous decisions of which store to purchase their durable and storable goods. Hence, a necessary condition for cross-channel substitution is that consumers visit multiple channels. Figure 3 illustrates that almost all consumers visit each channel with a probability greater than one. For instance, the lower left plot in Figure 3 depicts very little mass of consumers at both 0% and 100% visit probabilities for DG stores, while most of the mass is between 20% and 80% visit probabilities. Figure 4 focuses on the subset of consumers that bought blades on at least two occasions and depicts the proportion of blades bought at a given channel on the horizontal axis. We see that there is a significant mass of consumers that buy either all or none of their blades at Drug/Grocery and Mass Merchandiser stores. There is also a positive mass of consumers buying blades in these channels with some probability between 0 and 1. At Club stores, we see that no consumers who purchased at least twice bought all blades at club stores. Many consumers never bought blades at club stores but there is a mass of consumers that bought blades at all three formats. Together, Figure 3 and Figure 4 illustrate that customers visit multiple channels and exhibit some switching of the channel at which they buy blades.

We now present the results from the estimation of the model on these data.

⁹ A related concern here may be that the set of products available at a retail store is itself endogenous, enabling the manufacturer to price discriminate across retail formats. Modeling the endogenous selection of products at the retail level to address this issue is beyond the scope of the current analysis. Further, across the year of weekly data that we have, we do not see any variation in the set of products available at the outlets in our data. Hence, our approach is to condition on retail product availability in our analysis.

4 Model Estimates

The goal of our analysis is to estimate the model defined in Section 2 above with the data to conduct counterfactual changes in long-run price distributions to measure substitutability and/or complementarities between goods and retail formats. While our structural approach is necessary for analysis of these scenarios that are unobserved in the data, we also estimate several alternative (less structural) models to assess the fit of our model. These models include a static logit model without unobserved heterogeneity, a myopic version of our model with discount factor set = 0, a static “reduced form” model with unobserved heterogeneity, and with functions of past quantities and interpurchase times as covariates, the dynamic structural model without the state dependence term, as well as our full model. In general, we find that our preferred full model provide the best fit to the data, and also beat the other models on the Bayesian Information Criteria, which penalizes parameter proliferation. We find that accounting for unobserved heterogeneity as well as allowing for dependence of current purchases on past purchase history are both important to fit the patterns observed in the data. We also conduct extensive tests to assess the ability of our model to fit the observed patterns of the data (brand and format shares) both in and out of sample. We find that the fit is extremely good. Due to space constraints, we point the reader to Appendix C in the online appendix for these results.

Table 3 reports the estimates of our structural model. The brand utilities represent the value per-period of consumption of blades compatible with each razor technology (note that an exponent of the reported value must be taken to obtain the actual brand value). The estimates indicate that the mean utility from shaving with Mach is greater than shaving with other types of blades. The depreciation values for non-disposable and disposable razors can be calculated by taking the exponents of the reported values and dividing by one plus this exponent. They are nearly identical at 0.07 and 0.08 respectively. Thus, for the average consumer, blades typically last one or two weeks.

As described in the model above, we also estimate a parameter which helps quantify the resale value of old blades when adopting a new type of blade inventory. The mean of this parameter is about 0.5% (transforming -5.256 by its exponent divided by one plus its exponent). Even after considering the heterogeneity in this parameter, most of the customers’ salvage values are 3% or less. It seems that buying new blade types while still having old blades in stock is not a substantial factor in the decision process for these data. While not quantitatively significant, we retain this parameter because it is a useful modeling aspect that could be considered in applications with other datasets.

The price coefficient is 1.372 (i.e., $\exp(0.317)$) for the average consumer. We also find evidence of significant heterogeneity in the price effects in the population. This is a primary reason we estimate this parameter to be log-normally distributed. When estimating it to be normally distributed, we found it to be significant (with the proper sign), but also found that imposing a symmetric heterogeneity distribution led to some consumers in the tail with an improperly signed price coefficient. These price effects, along with the outlet visit probabilities will be primary determinants of the cross-store price elasticities.

--- **Tables 3 & 4 here** ---

The retail format visit parameters represent the average probability of visiting each of the formats (drug/grocery, mass merchandisers and club stores). We can see that consumers are estimated to visit the drug/grocery channel most frequently, with the mass merchandisers and then club stores following. The negative parameter on club store visit simply indicates that no-visit is more frequent than club store visits. The estimated visit probabilities are informative about the degree of retail market power over consumers. Grocery stores, which are visited most often, may have more market power over consumers since a consumer facing a price increase in that channel today anticipates that he has a high likelihood of returning to the same channel tomorrow, and facing a similar high price. This makes him more likely to buy at that high price, *ceteris paribus*. The high market power of grocery stores is also partly reflected in Table 2, where average prices per blade in those channels were higher. We also see that there is significant positive state dependence in the channel choice decisions. The standard deviations in these parameters indicate significant unobserved heterogeneity in channel visit probabilities.

Correlations Between Parameters

The correlations represent an important determinant of the substitutability between retail formats. In particular, the format visit probabilities could reflect the distribution of customers that are very loyal to each channel, or customers that are very similar and visit all channels with some probability. Strong positive correlations between channel utilities therefore indicate more inter-channel substitution, while negative correlations indicate little inter-channel substitution. These correlations are presented in Table 4. We see that mass merchandising channels are substitutable with both drug/grocery and club stores, based on positive correlations in preferences of 0.59 and 0.38 respectively. Yet, drug/grocery and club stores are not very substitutable with one another, indicated by a correlation of 0.09. Drug/grocery and club stores may therefore serve different segments of consumers. The estimated correlation in outlet preferences with price sensitivity also suggest that club stores may be attracting slightly less price sensitive consumers.

Many of the other correlations between heterogeneous parameters were estimated, but are not reported here.

Implications of Model Parameters

To get some intuition for what the model implies about purchase dynamics, we present plots of the value functions computed at the estimated parameter values. For illustration purposes, we pick a customer that uses about one Mach blade every two weeks (i.e., his net consumption is 26 blades in a year). We consider a situation when this consumer currently owns a Mach razor, and the currently owned blade is completely dull (i.e., $z = 0$). In the three panels of Figure 6, we present plots of the value of purchasing a blade pack to this consumer relative to the value of no-purchase at drug/grocery, mass merchandiser, and club stores respectively. In each of the figures, the horizontal axis is the number of unused blades in stock, b , and the vertical axis is the dollar value of surplus obtained when choosing the specified pack size relative to buying nothing. Looking at Figure 6, we see that the model implies that the more unused blades a consumer has, the less likely he is to buy any blades.

We now discuss how the model generates differing incentives for the same customer to purchase products at the different retail channels. To make the discussion concrete, we compare the incentive of this customer to purchase a Mach blade pack at each of the 3 outlets when owning 0 blades. Looking at the first panel, we see that the customer faces about a \$4 disincentive to purchase either of the Mach blade packs at a drug/grocery store (about \$2.50 of this is due to the purchase disutility). But for a large unobserved purchase incentive arising through a logit error, the customer will on average not purchase. A similar pattern holds at the mass merchandiser stores; the disincentive to purchase a Mach 4 pack is about \$3.50, implying that the customer will not refill at the mass merchandiser on average. However, looking at the third panel for club stores, we observe that the customer obtains a positive surplus when buying Mach refill blades at a club store. With zero unused Mach blades in inventory, this surplus is about \$0.26 for a Mach 16 pack and \$4.88 for a Mach 20 pack. Even when the customer has 8 Mach blades, he still realizes a surplus of about \$1.25 when refilling with a 20 pack at a club store. Given the large pack sizes available at the club store, and the low prices, the model implies that the customer at this state will strongly prefer to buy Mach blades only at Club stores. On average, the model will predict that the customer would want to wait when at a drug, grocery or mass merchandising store, so that he can refill his blades during a future visit to a club store, where per-blade prices are lowest.

Discussion

To summarize the demand estimates, we find strong evidence for heterogeneity in the data, as well as preferences for channel visits that are correlated with consumer tastes and consumption rates. Furthermore, the above discussion illustrates that the estimated model captures incentives for customers to optimally choose which store to purchase their refill blades from despite the fact that stores do not provide any direct utility to the consumer. Stores serve as a vehicle to obtain low prices if the customer is likely to visit. Our model illustrates that club stores, which specialize in selling large pack sizes at low per-unit prices, increase the incentive to stockpile and decrease incentives to purchase at other stores particularly for high volume customers like the biweekly blade user considered above.

5 Substitution Patterns for Tied Goods Sold Through Competing Retail Formats

In this section, we used the estimated model to measure a) complementarities between tied goods, and b) substitution patterns between shaving alternatives and retail formats. We measure each of these by simulating long-run own and cross-price elasticities across products and across retail formats. By *long-run*, we mean elasticities that incorporate the demand effects due to a change in current prices, as well as the substitution across stores and time due to changes in the price expectations of consumers. We measure these long-run effects by simulating consumer demand in response to a permanent change in the price distribution (e.g. EIK). In Appendix D, we provide a description of the procedure to simulate these values using the dynamic structural model we develop.

Complementarities Between Razors and Blades

We are interested in measuring complementarities between razors and blades as these determine the relative pricing of the primary and aftermarket goods for the manufacturer. The theory of pricing for tied goods suggests that under most plausible scenarios, the optimal strategy for the manufacturer is to lower margins on the primary good, and make up margins on the aftermarket good. This “razor-razorblades” pricing model is motivated by two arguments. The first flows from the existence of switching costs that partially prevent locked-in consumers from switching away from the system post their purchase of the primary good, thereby enabling the firm to extract rents through the sales of the aftermarket good. The second motivation arises from the potential for second degree price discrimination. By pricing the primary good low, the seller can attract low willingness to pay customers that demand few aftermarket goods. Pricing the aftermarket good higher then allows the seller to extract more surplus from higher willingness to pay customers

who consume more of the aftermarket goods. While we cannot sort between these two explanations, we will use our estimates to informally assess the empirical significance of either considerations.

A priori, economic switching costs in this category are likely to be low, since they should be bounded by the price of buying a new razor, which is low (the razor component of a Mach razor pack sold at Drug and Grocery or Mass Merchandisers is something less than \$1.95 and \$1.69 respectively). A psychological switching cost, related to habit persistence can however exist. Our estimates from Table 3 reveal that the switching cost from habit persistence is of the order of \$0.11, which we believe is not quantitatively large enough to support the switching cost explanation.

We now consider the long run elasticity of demand of Mach razors and blades with respect to manufacturer prices. We estimate these by simulating a permanent 1% price change by *all* retailers in the market. From the bottom two panels of Table 5, we see that if prices of Mach razor packs were increased by 1%, overall manufacturer derived demand from the market would decrease by 0.84%. A 1% permanent reduction in Mach blade prices on the other hand reduces Mach blade demand by 2.85%. Thus, our estimates reveal that Gillette is pricing Mach razors in the inelastic region of the demand curve, and pricing blades on the elastic region, which is consistent with the price discrimination incentive. While this is not a test of price discrimination per se, in conjunction with the low estimated switching costs, this suggests that discrimination is likely one of the primary forces driving pricing in this industry. Table 5 also finds evidence for significant complementarities between razors and blades in this industry.

Substitution between Competing Shaving Systems

The elasticities are also informative of substitution amongst competing systems. Given that Gillette sells multiple generations of tied good technologies, we explore how the pricing of their flagship tied good technology, the Mach, affects the demand of their previous generation technology, the Sensor, relative to demand for competing alternatives (Schick and Disposables). Looking at the bottom two panels of Table 5, we see that is evidence that increases in price for Mach blades results in substitution to blades of other technologies. A cursory reading of the elasticities suggests that substitution to Schick blades is the largest. This is however misleading since the Schick market share is quite small. In a separate calculation (not reported), we find that in terms of unit sales increases, about 83% of the increase in competing system blade sales induced by the 1% increase in Mach blade prices, consists of Sensor blades. This suggests that Gillette's pricing is influenced more by its decision to segment customers between its flagship technology and previous generations than it is by pressures from competing manufacturers.

Taken together, these elasticity results suggest that during this time period, Gillette may be best characterized as a monopolistic seller of razor systems that price discriminates by introducing newer generations targeted at the highest value customers. While analyzing product introductions is beyond the scope of this paper, it suggests that this is an interesting area of research for this market and perhaps other tied good systems.

Retail Substitution Patterns

We now use the model to study retail substitution effects in our data. We are primarily interested in these as they are informative of the extent to which downstream retail competition may distort the desired pricing of the manufacturer. In particular, presence of an oligopolistic retail channel introduces new avenues for channel conflict that are not present in non-tied product markets. Fundamentally, tying is difficult to sustain in the downstream. While the manufacturer is able to enforce a tie between the goods via, say product design, the retailer has trouble enforcing a tie since consumers can buy the aftermarket goods elsewhere. In the presence of retail market power, this aspect creates two problems for the manufacturer. The first, double marginalization, is exacerbated by the tied goods nature of the products because higher prices reduce adoption of the primary good either directly or indirectly through complementary aftermarket goods. The second issue, which we call “cross-product horizontal externalities,” arises when one store’s price increase reduces the demand of complementary goods sold at another store. While many tied good products are sold through retail channels, surprisingly, to our knowledge, both the theory literature and the empirical work on tied goods so far, has ignored the role of downstream retail competition we consider. While exploring a dynamic channel game formally is beyond the scope of this paper, by empirically analyzing substitution patterns at the retail level, we hope to shed light on these channel problems and motivate future work modeling the supply side implications of selling tied goods through oligopolistic downstream channels.

Double marginalization is particularly problematic for tied goods because the manufacturer requires a low price for the primary good to serve as a customer acquisition device to drive a flow of profitable aftermarket sales. A retailer with sufficient market power to raise razor prices could harm the manufacturer by generating a demand contraction in the razor market that cuts off all of the manufacturer’s profitable blade sales associated with the lost razor sales. This is plausible in the razor and blades market we consider because one downstream channel (drug/grocery stores) sells a large number of razors, and may have significant market power. Retail market power in the aftermarket is problematic as well, since this implies “too high” blade prices from the manufacturers’ point of view. This results in reduced blade output and possibly a contraction in razor sales due to the reduced demand for razors from forward-looking

consumers who anticipate high blade prices following purchase. Consequently, tied good manufacturers, such as Gillette in the razor blade market, desire retail competition in *both* the razor market and the blade market.^{10,11}

In addition, complementarities in demand between the primary good sold at one retail format, and the aftermarket good sold at a different retail format creates a “cross-product horizontal externalities” problem. The term *horizontal externalities* has generally been used to describe channel conflicts such as classic dealer free-riding, in which a discount chain sells a good at a low price with low sales support to a customer who learned about the good at a high-price, high-support channel (see Mathewson and Winter, 1986; Klein and Murphy, 1988; Iyer, 1988). If sales support or service is not desired by the manufacturer, these externalities typically do not exist. However, in the case of tied goods, a “cross-product horizontal externality” can still arise if the pricing decisions of a retailer specializing in the aftermarket good significantly affects the demand at a different retailer specializing in the primary good, or vice-versa. For example, if a retailer specializing in blades raises prices and consequently reduces aggregate blade demand, it will do so without internalizing the lost razor sales at other formats. Significant cross-format complementarities of this nature arise only in the presence of retail market-power. For instance, in the above example, without retail pricing-power in blades, a consumer’s razor purchase decision at a given format is unaffected by blade price changes at another, since he can easily switch to a different retail format for his blade requirements. Razors and blades provide a particularly interesting example because, in the data, club stores primarily sell blades, while Drug and Grocery stores sell many more razors relative to blades (The ratio of Mach blades to razors sold in the data by format are: 47 drug/grocery, 74 Mass Merch. and 221, Club). To evaluate the extent of “cross-product horizontal externalities,” researchers must therefore assess whether any one channel can reduce aggregate demand of either the primary or aftermarket good

¹⁰ Competition in either one of the markets does not eliminate the distortion. Suppose the tying good is sold through a perfectly competitive retail channel, but the tied good is sold through retailers with market power. Perfect competition in the tying good enables the manufacturer to pass through desired low razor-prices, but, the presence of market power in the blade channel implies that blade prices will be too high, and blade-output too low from his perspective. Given higher blade prices, razor demand will also be too low. The vertical externality persists if the tied good retail channel is perfectly competitive, but retailers selling the tying good have market power. The manufacturer would be able to pass-through his desired high price on blades, but double marginalization in the razor market imply that razor prices will be too high from his perspective. Overall demand for razors, and thus of blades, is again too low.

¹¹ Even if there is significant retail market power, the manufacturer can avoid double marginalization by using vertical restraints (e.g. resale price maintenance) and/or two-part tariffs (Moorthy 1987; Desai and Srinivasan 1995; Villas-Boas 1998; see Iyer and Padmanabhan 2003 for a review of the theory and Villas-Boas 2007 and Lafontaine and Slade 2005 for a review of the empirical work). Evaluating the incidence or efficacy of these vertical restraints is beyond the scope of the current analysis and the available data (we do not observe wholesale prices or details of the contract between the manufacturer and the retailer). Rather, our analysis here discusses whether there exists sufficient downstream retail substitution such that this may or may not be a quantitatively significant issue for the manufacturer.

and, if so, whether the reduction is large enough to significantly impact other channels' demands for the complementary good. If the externality exists, channels will not internalize all of the negative effects of increasing prices and will therefore likely reduce adoption of the tied good system.

We now discuss how the elasticity estimates are informative of both issues.

Retail Pricing Power

Given the lack of a formal model of pricing, we cannot assess equilibrium price responses in the downstream channel game. However, the demand model can still be used to assess substitution patterns, which form an important driver of pricing power at the retail level. Our approach is to simulate how much of the manufacturers' *overall razors* sales would be lost if one of the retail formats unilaterally increased its *razor* price. If consumers simply switch to other stores in response to the price increase, the overall reduction in sales to the manufacturer will be small, and we may infer that retail substitution is strong, and that the channel's pricing power is small. This will imply that the potential for double marginalization is low. We focus on simulation on Drug and Grocery stores since they are the primary sellers of razors in the data. Looking at Table 5, if DG stores raised Mach razor pack prices by 1%, they would lose 2.25% of their Mach razor demand. The cross-price effects of 1.21 and 0.55 for Mass Merchandisers and Club stores respectively, indicate significant substitution within Mach to other retail formats. Put differently, in terms of unit sales, this implies that 67% of Mach razor sales that would have been lost at the Drug/Grocery channel would be made up at the competing retail formats. Overall, the price increase results in only a 0.31% decrease in the total Mach razors sold by Gillette across all retail formats in this market. We interpret these results as suggesting that while the Drug/Grocery channel does have some degree of pricing power, retail format competition does guard Gillette from a concern that every razor sale lost at the Drug/Grocery channel is permanently lost.

Table 5 also reveals that long-term demand for Mach blades is highly elastic at both the Drug/Grocery format (own-price elasticity of -4.20), and at Club format stores (own-price elasticity of -3.13). While there is evidence of retail substitution in blades in response to price changes at both, Club store pricing imposes the largest impact on the manufacturer, since a 1% increase in Mach blade prices at Club stores can decrease the manufacturer's Mach blade sales by 1.28%. In terms of unit sales, this implies that only 19% of the Mach blade sales lost at a club store would be shifted over to competing retail formats (the rest of the 1.28% reduction in blade demand is permanently lost to Mach). The larger effect of Club stores are likely driven by the larger pack sizes they sell, as well as their patronage of the highest usage Mach customers. The estimates

suggest there is some potential for Club store pricing power in blades, depending on the mode of competition.

Cross-Product Horizontal Externalities

We now assess the potential for cross-product horizontal externalities. In particular, we test whether price increases in Mach razors sold at one format significantly affects demand for Mach blades sold at another, and vice versa. We begin by assessing the affects of razor price increases at a Drug/Grocery format stores. An increase in razor prices at this format can reduce blade demand at other formats by a) delaying the purchase of the razor and hence reducing blade consumption, or b) by eliminating the entire stream of blade sales associated with the Mach razors that may never be purchased. Referring to the top panel of Table 5, we see (a) a 1% price increase in Mach razors at the Drug/Grocery channel has a negligible effect on the sales of blades at the mass merchandisers and club stores (cross-price effects of -0.08 and -0.06); and (b) aggregate blade demand to the manufacturer is also only slightly affected (reduction of 0.08%). Simulations of the model (not reported) indicate that this is driven by the fact that marginal customers that no longer adopt razors are lower than average users of blades. Blade price increases at Drug/Grocery stores, however, seem have more significant effects on overall demand (reduction of 0.31%), and demand for razors at other stores. Table 5 also reports on the effects of raising blade prices at Club stores. We see that a 1% increase in blade prices at club stores would reduce razor adoption by only 0.11% at Drug and Grocery stores and 0.08% at mass merchandisers. Overall manufacturer razor demand is only reduced by 0.08%. These small effects suggest that the effects of blade price increases are primarily to reduce blade usage by Mach customers, as opposed to reducing the number of Mach razor adopters.

Take away

These long-run effects suggest that the downstream retail channel for this industry, while oligopolistic, involves substitution between retail formats. From the manufacturer's perspective, it seems the greatest potential area of concern for pricing power at the retail level might be in blades sold by club stores and drug and grocery stores. The club format represents about 18% of Mach blade demand, and unlike drug and grocery, is also a concentrated retail format. Reputational incentives for maintaining a low-price image may however be a countervailing force that prevent club stores from exploiting potential market power in pricing. While our estimates suggest some pricing power at the format level amongst Drug/Grocery stores, the prevalence of a large number of Drug/Grocery stores also suggests additional within-format, inter-store competition that may reduce the exercise of this pricing power. Overall, our results therefore

suggest that retail competition in this market is such that manufacturers may be able to pass much of their desired pricing through to the downstream channel without the need for additional vertical restraints.

6 Conclusions

This paper makes three contributions. First, we develop a dynamic structural demand model for tied goods sold through competing retail formats. Second, we empirically measure retail format substitution in a real-world market context, and discuss the potential pricing implications for primary and aftermarket goods through an oligopoly retail environment. Third, we identify the horizontal externality aspect of retail competition in tied goods, which has hitherto not been articulated in the literature.

The demand model we propose advances the literature on tied goods and on estimation of demand for durable and storable goods. While modeling tied goods necessitates a model that allows the choice set to be conditional on past purchases (i.e., past purchase of a primary good), the fact that many aftermarket goods are purchased and consumed in subsequent periods requires a forward-looking model that could allow demand for the primary good to be impacted by expected future prices over the aftermarket goods. Furthermore, the large pack sizes that blades are sold in requires an underlying model of how customers stock blades in inventory, and then endogenously replace them over time.

While these characteristics are necessary for any dynamic tied goods context, they can be generalized to other storable or durable good contexts. For instance, modeling a consumer to be able to purchase from the available products at the store currently visited or to wait and buy at the same or a different channel in the future can help measure retail substitution in many single category demand analyses without a tying aspect. In this sense, we add to the relatively small empirical literature in marketing on measuring store substitution. While we find little evidence for complementarities in demand between razors sold at one store and blades sold at another, we find significant substitutability in razors sold across retail formats. The latter is in contrast to much of the past literature, which has found little or small across store-price effects. We speculate that much of cross-store substitution is driven by intertemporal-switching, which would be missed by the past literature, which has primarily worked with static models in this context. While these intertemporal effects are accounted for in our measure of long-run effects, these would be missing from static short-run estimates of store substitution patterns.

References

- Akerberg, D., (2001) "A New Use of Importance Sampling to Reduce Computational Burden in Simulation Estimation" working paper.
- Bell, David R. and James M. Lattin (1998) "Shopping Behavior and Consumer Preference for Retail Price Format: Why 'Large Basket' Shoppers Prefer EDLP," *Marketing Science*, 17, 1 (Winter), 66-88.
- Borenstein, S., J.K. MacKie-Mason and J. Netz, (1994) "The Economics of Customer Lock-In and Market Power in Services," in Harker, Patrick T. (ed.) *The Service Productivity and Quality Challenge*. Kluwer Academic Press.
- Bronnenberg, B, Kruger, M. and Mela, C. (2008), "The IRI Marketing Data Set," *Marketing Science*, 27 (4), pp. 745-748.
- Borenstein, S., J.K. MacKie-Mason and J. Netz, (2000), "Exercising Market Power in Proprietary Aftermarkets." *Journal of Economics and Management Strategy*, 9(2), 157-188.
- Burstein, M. L. (1960) "The Economics of Tie-In Sales," *Review of Economics and Statistics*, Vol. 42, No. 1. (Feb) pp. 68-73.
- Carlton, D. and Waldman, M. "Theories of Tying and Implications For Antitrust," in W. Collins ed. *Issues in Competition Law and Policy*, forthcoming.
- Desai, Preyas and Kannan Srinivasan (1995), "Demand Signaling Under Unobservable Effort in Franchising: Linear and Nonlinear Contracts," *Management Science*, 41 (10), 1608-1623.
- Dubé, Jean-Pierre, Günter J. Hitsch and Pradeep Chintagunta (2007). "Tipping and Concentration in Markets with Indirect Network Effects," working paper, University of Chicago.
- Dubé, Jean-Pierre, Günter J. Hitsch and Peter Rossi (2008). "Do Switching Costs Make Markets Less Competitive?," *Journal of Marketing Research*, forthcoming.
- Erdem, Tülin, Susumu Imai and Michael Keane (2003), "A Model of Consumer Brand and Quantity Choice Dynamics under Price Uncertainty," *Quantitative Marketing and Economics*, 1 (1), 5-64.
- Gil, R. and W.R. Hartmann (2007) "Metering Price Discrimination: An Empirical Analysis of Concession Sales at Movie Theaters," working paper, Stanford University.
- Gauri, D., K. Sudhir, and D. Talukdar (2007). "The Temporal and Spatial Dimensions of Price Search: Insights from Matching Household Survey and Purchase Data" *Journal of Marketing Research*.
- Gordon, Brett (2007), "A Dynamic Model of Consumer Replacement Cycles in the PC Processor Industry," working paper, Columbia University.
- Hall, R. (1997). "The Inkjet Aftermarket: An Economic Analysis", working paper, Stanford University.
- Heckman, J. (1981). "Heterogeneity and State Dependence," in S. Rosen (ed.), *Studies in Labor Markets*, University of Chicago Press, 91-139.
- Hendel, I. and Aviv Nevo (2006) "Measuring the Implications of Sales and Consumer Inventory Behavior," *Econometrica*, November 74(6), 1637-73.
- Iyer, Ganesh (1988) "Coordinating Channels Under Price and Nonprice Competition," *Marketing Science*, 17(4), 338-355.
- Iyer, Ganesh. and V. Padmanabhan (2003), "Contractual Relationships and Coordination in Distribution Channels," in *Managing Business Interfaces: Marketing, Engineering, and Manufacturing Perspectives*, ed. A. Chakravarty and J. Eliashberg, Kluwer Academic Press, 2003.
- Klein, B. (1996). "Market Power in Aftermarkets," *Managerial and Decision Economics*, 17, 143-164.
- Klein, B. and K.M. Murphy (1988). "Vertical Restraints as Contract Enforcement Mechanisms," *Journal of Law and Economics*, XXXI, 265-297
- Lafontaine, Francine and Margaret Slade (2005), "Exclusive Contracts and Vertical Restraints: Empirical Evidence and Public Policy," forthcoming in *Handbook of Antitrust Economics*, Paolo Buccirossi (ed.) Cambridge: MIT Press
- Lee, Robin (2007). "Vertical Integration and Exclusivity in Platform and Two-Sided Markets", working paper, Harvard University.

Mathewson, Frank and Ralph Winter. (1986). "The Economics of Vertical Restraints in Distribution," in *New Developments in the Analysis of Market Structures*, ed. Frank Mathewson and Joseph Stiglitz, Cambridge, MA. MIT Press.

Moorthy, K. S. (1987). "Managing Channel Profits: Comment," *Marketing Science*, 6, 375-379.

Nair, Harikesh, Pradeep Chintagunta and Jean-Pierre Dubé. (2004). "Empirical Analysis of Indirect Network Effects in the Market for Personal Digital Assistants," *Quantitative Marketing and Economics* 2(1), 23-58.

Oi, W.Y. (1971) "A Disneyland Dilemma: Two-part Tariffs for a Mickey Mouse Monopoly," *Quarterly Journal of Economics* 85, pp. 77-96.

Roy, R. Chintagunta, P. and Halder, S. (1996) "A Framework for Analyzing Habits, 'Hand-of-Past,' and Heterogeneity in Dynamic Brand Choice," *Marketing Science*, 15, 3, 280-299.

Rust, J. (1987). "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica* 55(5), 999-1033.

Rust, J. (1996). "Numerical Dynamic Programming in Economics", in Amman, H., Kendrick, D. and Rust, J. (eds.), *Handbook of Computational Economics*, Elsevier, North Holland.

Schmalensee, R. (1981). "Monopolistic Two-Part Pricing Arrangements", *Bell Journal of Economics*, 8, 445-467.

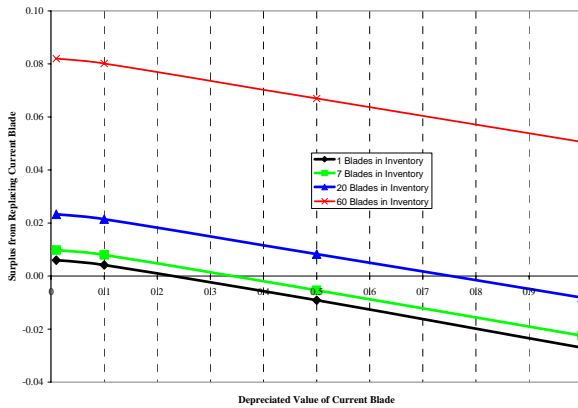
Shapiro, Carl, (1995) "Aftermarkets and Consumer Welfare: Making Sense of Kodak." *Antitrust Law Journal* 63(2), Winter, 483 - 4511.

Sun, B. (2005), "Promotion Effect on Endogenous Consumption," *Marketing Science* 24(3), pp. 430-443.

Villas-Boas, Miguel (1998) "Product Line Design for a Distribution Channel," *Marketing Science*, 17 (2), pp. 156-169.

Villas-Boas, Sofia-Berto (2007) "Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data," *Review of Economic Studies*, 74(2), 625-652.

Figure 1: Endogenous Blade Replacement



Notes: The plots present the difference in surplus from using a new blade next period versus continuing to use the current blade, as a function of the depreciated value of the current blade (z). Each line represents a different value for the inventory of unused blades b . The plots are for a consumer currently owning a Mach razor and blades.

Figure 2: Distribution of Blade Usage in Data

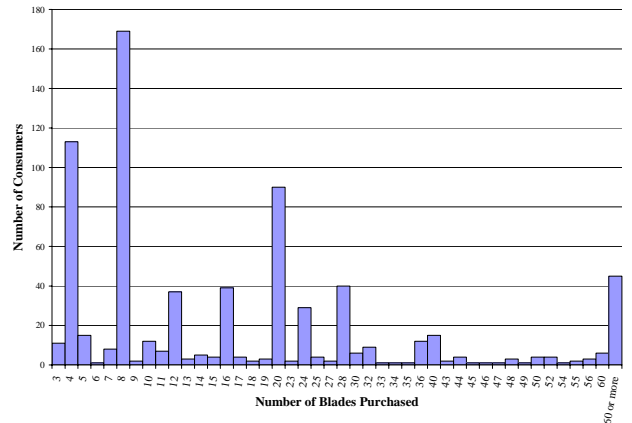
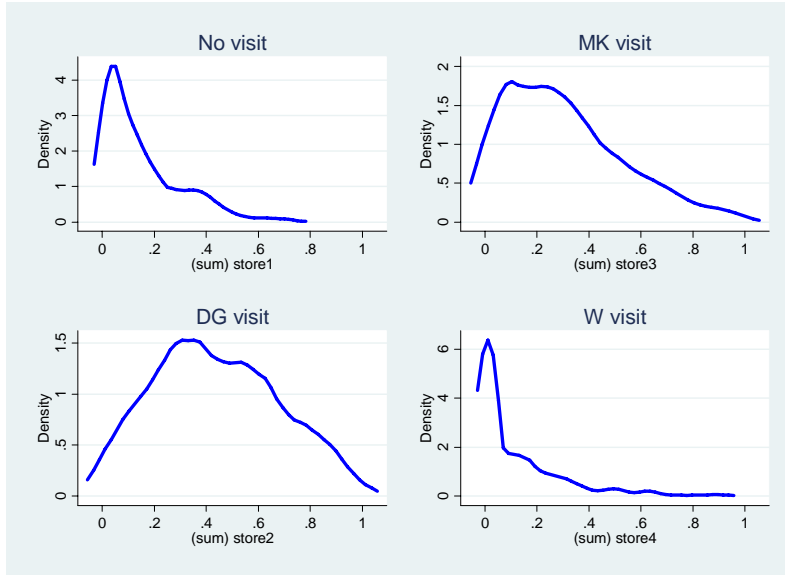
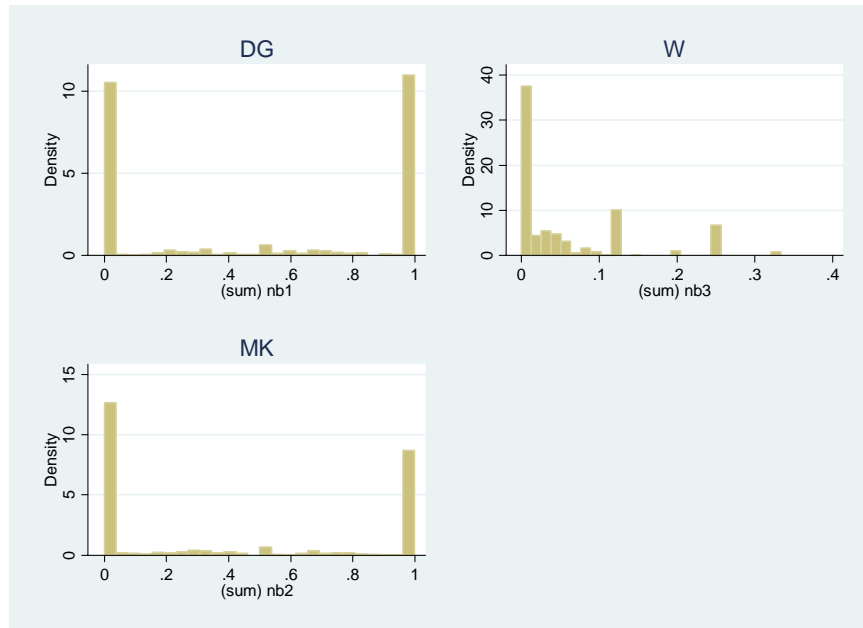


Figure 3: Distribution of Store Visits Across Consumers



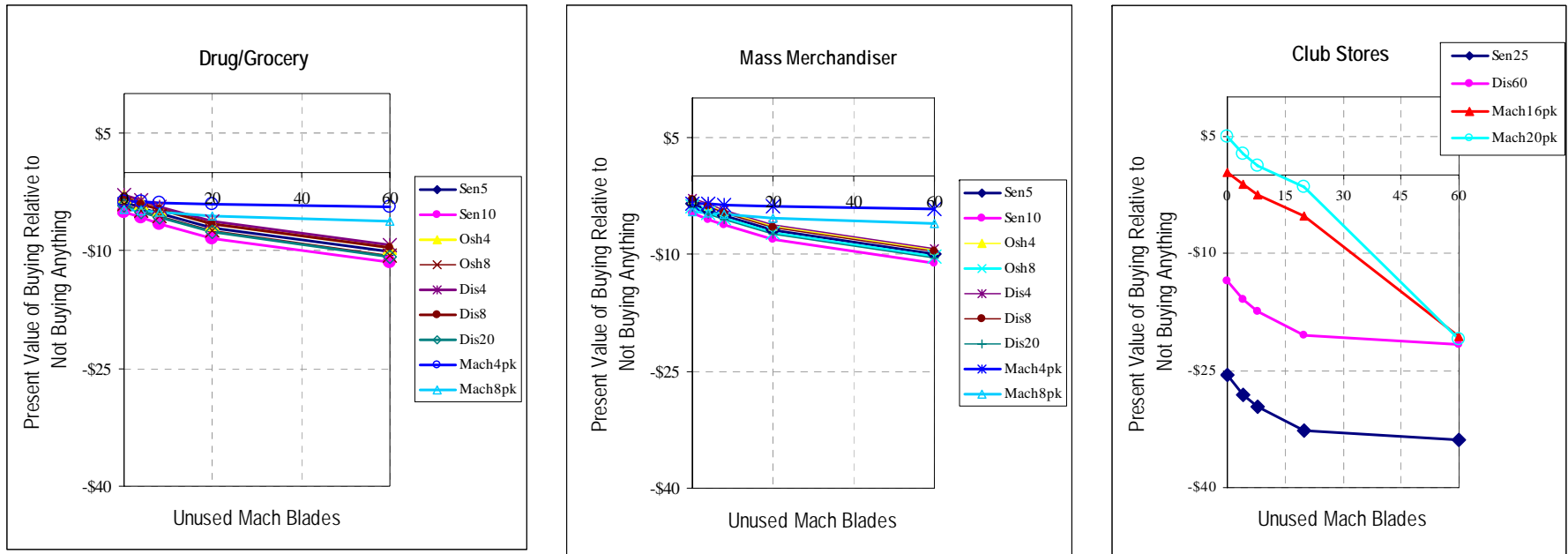
Notes: This plot shows the distribution of store-visits by consumers. The plot is generated in the following way: for each consumer, we calculate the proportion of weeks in which either no store was visited, or one of the DG, MK or W stores were visited. The figure plots the density of these proportions across consumers. For example, the first plot (top left) shows that most consumers do not have a large proportion of weeks in which no-store is visited; the plots on the top-right and bottom-left show that many consumers are likely to make a large proportion of DG and MK visits; the bottom right panel shows that across most consumers, W store visits is liable to be <10% of total visits

Figure 4: Proportion of Blade Purchases by Channel



Notes: This plot shows the distribution of the proportion of blade purchases made by consumers across channels. For each consumer, we calculate the proportion of blades bought at one of the DG, MK or W stores. The figure plots the density of these proportions across consumers. We see that there are many consumers that have made all their blade purchases at a DG or an MK store (significant mass at 1 in the panels on the left). At the same time, we also see that there are many consumers who have shopped for blades across the three channels.

Figure 6. Relative value of buying a pack at each retail outlet for a biweekly Mach blade user currently owning a Mach razor and blades



Notes: The plots present the value of buying a pack relative to the value of no-purchase as a function of the current blade stock b , evaluated at $z=0$, at each retail outlet, for a consumer currently owning a Mach razor and blades, who replaces his blades on a biweekly basis.

Table 1
Summary Statistics
Store Visits and Razor Technology Ownership

	Obs	Mean	Std. Dev.	Min	Max
Store Visits					
No Store	40,600	0.15	0.35	0	1
Grocery or Drug	40,600	0.44	0.50	0	1
Mass Merchandiser	40,600	0.30	0.46	0	1
Club Store	40,600	0.11	0.31	0	1
Mach Razor Owned	40,600	0.31	0.46	0	1
Blade Type In Consumer Inventory					
Gillette - Sensor	40,600	0.09	0.28	0	1
Gillette - Mach	40,600	0.29	0.45	0	1
Schick	40,600	0.04	0.21	0	1
Disposable	40,600	0.58	0.49	0	1

Table 2
Pack Size Average Prices by Razor Technology and Store Type

	Drug and Grocery		
	Stores	Mass Merchandisers	Club Stores
Gillette - Mach			
3pk Razor	\$8.10 (\$2.70)	\$7.36 (\$2.45)	
7pk Razor			\$14.70 (\$2.10)
4pk Blades	\$8.20 (\$2.05)	\$7.57 (\$1.89)	
8pk Blades	\$15.51 (\$1.94)	\$14.42 (\$1.80)	
16pk Blades			\$26.33 (\$1.65)
20pk Blades			\$28.03 (\$1.40)
Gillette - Sensor			
5pk Blades	\$7.13 (\$1.43)	\$6.42 (\$1.28)	
10pk Blades	\$13.31 (\$1.33)	\$12.06 (\$1.21)	
25pk Blades			\$25.46 (\$1.02)
Schick			
4pk Blades	\$6.57 (\$1.64)	\$6.26 (\$1.56)	
8pk Blades	\$13.13 (\$1.64)	\$12.51 (\$1.56)	
Disposables			
4pk Blades	\$2.12 (\$0.53)	\$1.92 (\$0.48)	
8pk Blades	\$4.25 (\$0.53)	\$3.85 (\$0.48)	
20pk Blades	\$10.62 (\$0.53)	\$9.62 (\$0.48)	
60pk Blades			\$16.18 (\$0.27)

Notes: Average price per unit in parenthesis

Table 3
Model Estimates

	Random Coefficients		
	Mean	Standard Deviation	
Brand Utilities			
Sensor	-1.271 (0.138)	1.016 (0.020)	a
Mach	-0.176 (0.074)	0.888 (0.020)	a
Schick	-2.015 (0.114)	1.423 (0.025)	a
Disposables	-1.317 (0.159)	2.349 (0.029)	a
Price Coefficient	0.317 (0.120)	1.787 (0.022)	a
Blade Depreciation Rate			
Non-Disposables	-2.577 (0.115)	1.506 (0.027)	b
Disposables	-2.377 (0.127)	1.881 (0.023)	b
Resale Percentage	-5.256 (0.144)	1.462 (0.035)	b
Habit Persistence	0.153 (0.007)	0.114 (0.005)	
Purchase Disutility	-3.483 (0.017)	0.238 (0.014)	
Store 1	1.435 (0.104)	1.464 (0.030)	
Store 2	0.713 (0.102)	1.658 (0.046)	
Store 3	-1.463 (0.332)	2.819 (0.085)	
Store Choice Persistence	0.095 (0.029)	0.233 (0.010)	

Notes on Table 3: standard errors in parenthesis. 'a' denotes that the variable is log-normally distributed. The coefficient for these variables can be obtained by taking the exponent of the reported value. 'b' denotes that the variable is constrained to be between 0 and 1. The coefficient for these variables can be obtained by taking the exponent of the reported value and dividing it by 1 plus this exponent.

Table 4
Correlations in Store Visit Probabilities and Price Sensitivity

	Store Choice: Grocery/Drug	Mass Merch.	Club Stores
Store Choice			
Grocery/Drug	1	0.59	0.09
Mass Merch.	0.59	1	0.38
Club Stores	0.09	0.38	1
Price Sensitivity	0.07	0.00	-0.06

Table 5
Long-Run Demand Elasticities
with Respect to Mach Razor and Blade Prices

	Retail Channels			Manufacturer
	Drug/Grocery	Mass Merchandiser	Club Stores	
<i>Drug/Grocery: Mach Razor Pack Price</i>				
Mach Razor Pack Demand	-2.25	1.21	0.55	-0.31
Mach Blade Refill Demand	-0.14	-0.08	-0.06	-0.08
Sensor Blade Demand	0.01	-0.01	0.00	0.00
Schick Blade Demand	0.27	-0.05		0.11
Disposable Blade Demand	0.26	-0.05	-0.02	0.10
<i>Drug/Grocery: Mach Blade Refill Price</i>				
Mach Razor Pack Demand	-0.45	-0.26	-0.07	-0.31
Mach Blade Refill Demand	-4.20	0.57	0.28	-0.70
Sensor Blade Demand	0.02	0.00	0.00	0.01
Schick Blade Demand	0.64	0.14	0.00	0.38
Disposable Blade Demand	0.20	0.00	0.81	0.10
<i>Club Stores: Mach Blade Refill Price</i>				
Mach Razor Pack Demand	-0.11	-0.08	0.08	-0.08
Mach Blade Refill Demand	0.58	0.67	-3.13	-1.28
Sensor Blade Demand	0.00	0.00	0.04	0.01
Schick Blade Demand	0.13	0.12	0.00	0.13
Disposable Blade Demand	0.11	0.08	1.80	0.11
<i>All Stores: Mach Razor Pack Price</i>				
Mach Razor Pack Demand	-0.88	-0.54	-1.77	-0.84
Mach Blade Refill Demand	-0.23	-0.19	-0.05	-0.13
Sensor Blade Demand	0.00	0.01	0.01	0.01
Schick Blade Demand	0.25	0.24	0.00	0.24
Disposable Blade Demand	0.22	-0.07	2.05	0.08
<i>All Stores: Mach Blade Refill Price</i>				
Mach Razor Pack Demand	-0.91	-0.74	-0.04	-0.72
Mach Blade Refill Demand	-3.53	-3.06	-2.42	-2.85
Sensor Blade Demand	0.02	0.03	0.03	0.02
Schick Blade Demand	0.95	0.73	0.00	0.84
Disposable Blade Demand	0.29	0.34	2.97	0.33

Notes: Table reports the % change in long run-demand due to a 1% change in row prices