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: December 4, 2007

Abstract  
We empirically investigate the demand for tied goods sold through competing retail channels. Tied good pricing strategies commonly involve a low price on the initial purchase (i.e. the primary good) to drive adoption, and a substantial markup on aftermarket goods to capture value. However, if the goods are sold through downstream channels, retail market power and a misalignment of incentives could distort the relative prices of primary and aftermarket goods. To evaluate whether retail competition is strong enough to prevent such distortions, we explore the commonly noted example of razors and blades, which are sold through drug, grocery, mass merchandising, and club stores. We specify a forward-looking demand model that incorporates dynamics arising from the tied good nature of the products and the stockpiling and durability aspects of razors and blades. Furthermore, we allow intertemporal substitution in the purchase of both razors and blades to occur across channels as well as time. This modeling feature enables a novel approach to measuring retail competition in single category demand analyses. Our estimates indicate that there is substantial cross-channel substitution in razors, but some retail market power in blades. However, the channel with the most market power in blades, club stores, specializes in high volume customers that would adopt a razor even if blade prices are higher. This suggests that the manufacturer can achieve its desired level of razor adoption without vertical restraints, though blade sales may be slightly reduced by double marginalization.

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

* We thank Michaela Draganska, J.P. Dube, Ganesh Iyer, Jim Lattin, Yesim Orhun, K. Sudhir, Miguel Villas-Boas as well as seminar participants at the 2007 Summer Institute in Competitive Strategy and Yale SOM for their helpful comments. We also thank Pradeep Chintagunta for his 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 empirically measure the extent of retail competition in the sales of tied goods. 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. Tie-in sales 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 theory of pricing for tied goods typically finds that, relative to separate component pricing, the manufacturer has an incentive to reduce prices on the primary product to drive adoption, while shifting margins to aftermarket products. ¹ This paper analyzes whether or not a manufacturer can obtain this desired pricing when selling the tied goods through retail channels.

The presence of a retail channel introduces the potential for channel conflict arising from misaligned incentives. 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 that distorts his desired pricing in the downstream channel. 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 problem, which we call “cross-product horizontal externalities,” arises when one store’s price increase reduces the demand of complementary goods sold at another store. The goal of this paper is to develop and estimate a demand model that measures downstream market power and retail competition in a specific tied goods context (razors and blades) to determine the extent to which either of these problems are of quantitative significance for the manufacturer. While many tied good products are sold through retail channels, surprisingly both the theory literature and the empirical work on tied goods has ignored the role of downstream retail competition.

Double marginalization (Spengler 1950; Tirole 1988) is particularly problematic for tied goods because the manufacturer requires a low price for the primary good to serve as a

¹ This enables the manufacturer to either extract higher margins on the aftermarket good from locked-in consumers (Borenstein, MacKie-Mason & Netz 1994, 2000) or to use the tie to price discriminate by metering. The latter motivation arises when consumer demand for the aftermarket product may help the firm “meter” their willingness to pay for the primary good. This has been explored by Burstein (1960), Oi (1971), Schmalensee (1981), Klein (1996) and recently evaluated empirically by Gil and Hartmann (2007).
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.²³

In addition, complementarities in demand between the primary good sold at one retail channel, and the aftermarket good sold at a different retail channel 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 Telser, 1960; 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 (primary) good affect the demand at a different retailer specializing in the primary (aftermarket). 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 channels. In other words, the cross-channel-cross-product price elasticities are

² 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.

³ 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 discusses whether there exists sufficient downstream retail competition such that this may or may not be a quantitatively significant issue for the manufacturer.
negative due to the complementarities between the products such that a price increase at one channel significantly reduces demand at other channels. Significant cross-channel complementarities of this nature arise only in the presence of retail market-power. For instance, without retail market-power in blades, a consumer’s razor purchase decision at a given channel is unaffected by blade price changes at another, since he can easily switch to a different retail channel for his blade requirements. Razors and blades provide a particularly interesting example because, as Figure 1 illustrates, club stores primarily sell blades, while Drug and Grocery stores sell many more razors relative to blades. 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 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.

From the discussion above, it is clear that measuring retail competition is of key relevance for tied goods. We develop a model that can be used to analyze retail competition and consumer demand in a tied goods market. We then estimate the model and use it to measure the cross-price elasticities relevant to channel competition in this industry. Measuring these price effects is complicated by the fact that razors and blades, like many other tied good products, are durable and storable, such that consumers may intertemporally substitute purchases. Others, such as Erdem, Imai and Keane (2003) and Hendel and Nevo (2006) have shown that dynamic demand models are useful for properly measuring elasticities in such settings. Following this work, we define a dynamic discrete choice model that accounts for the storability and durability of razors and blades. We build on the existing work by endogenizing per-period consumption, as well as the store at which customers accumulate inventory. The latter component is critical to measuring retail competition for durable or storable goods and constitutes a novel approach for measuring retail competition in single-category demand analyses.

While our goal is to evaluate whether retail competition is sufficient to ensure manufacturers can maintain their desired relative pricing between razors and blades and avoid horizontal externalities, rather than explain the source of the manufacturers desired relative pricing or decision to tie, it is instructive to point out features of this market that are consistent or inconsistent with existing theories of tying. A lock-in theory of high aftermarket prices is unlikely from a purely economic view of consumer behavior because the switching cost is approximately equal to the replacement cost of a razor, which is small. If there are however substantial psychological costs of switching razors, lock-in could explain high aftermarket prices. Firms may also tie blades to razors to engage in metering price discrimination which exploits a positive correlation in willingness to pay for razors and demand for blades. The fact that there is significant heterogeneity in blade demands suggests this is a viable explanation. Finally, the tie could merely be a product design necessity and manufacturers price razors low to facilitate trial. Throughout the remainder of the paper, we are agnostic about the source of the manufacturers pricing incentives.
Specifically, our model defines consumers’ decisions each period to buy razors and blades from the set of stores in the market. Forward-looking consumers incorporate the effect of their current choices on future payoffs, and take into account that the purchase of a razor today may lock them in to consuming blades associated with that razor technology in the future. The structural derivation also enables us to model the process by which a customer weighs buying a package of blades while on a trip to the grocery store against waiting until the next trip to a club store to buy the blades at lower prices. The propensity to wait is a function of consumer expectations about future prices, as well as expectations regarding the probability of their future store visits. This accommodates, for instance, the phenomenon that consumers may purchase and stockpile more blades during visits to club stores than grocery stores, since they anticipate during their club store visits, that their chance of visiting another club store in the immediate future is small.

In addition, we address the problem of modeling demand for tied goods in a fully dynamic, structural framework. One fundamental feature of the model is that the choice-set faced by the consumer (i.e., the set of blade packs to choose from) is dependent on his ownership of the primary good (i.e., the razor owned). A challenging aspect of this is the large number of pack sizes available and the discrete nature of purchased quantities. We accommodate this aspect of the market by deriving the demand system from discrete choices over aftermarket alternatives. This feature of the model could also be generalized 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 2005; Clements and Ohashi 2005; Dubé, Hitsch and Chintagunta 2007; Lee 2007). Unlike tied goods, those with indirect network effects exhibit demand side economies of scale that affect the quality, scope, and prices of aftermarket items (e.g., video games).

We apply the model to panel data on consumer purchases of razors and blades in a large Midwestern city in the US. The data span one year between 2002 and 2003. During this time, the market leader, Gillette, marketed the Mach as its flagship razor technology, though they still sold a substantial number of refill blades for the previous generation technology, the Sensor. The other competitor, Schick, had not yet introduced the Quattro and was marketing a couple of other technologies with little success. We observe all purchases of these two brands and disposable razor blades for every household across all stores, including wholesale and club stores. This is important since it enables us to construct the evolution of the stock of blades in the consumer's inventory, which is a key state variable driving the dynamics of consumer
demand. Furthermore, this allows us to infer substitution patterns across all channels in the market (grocery and drug stores, mass merchandisers, and wholesale club stores). Our results reveal that there is sufficient retail competition to prevent a channel member with misaligned incentives from reducing razor adoption below the level desired by the manufacturer. Specifically, we find cross-store price effects in the razor market suggestive of significant retail competition. A 1% increase in long-run price distribution for razors at the drug or grocery channel, which should have the greatest market power, creates significant substitution to other channels such that overall manufacturer demand for razors is only reduced by 0.01%. Razor adoption could also be reduced if a channel could raise blade prices to consumers on the margin of buying a razor. We find some evidence of market power in blades, but it has a negligible effect on razor demand because retailers with market power in blades either specialize in selling to high volume customers that would adopt the razor anyway or actually internalize the effect of blade prices on their own razor demand. These findings therefore assure manufacturers that there is sufficient inter-channel retail competition to obtain their desired level of razor adoption.  

To summarize, the paper makes the following contributions. From a modeling perspective, we extend storable good dynamic discrete choice models to allow consumers to strategically select the store at which they purchase durable goods and accumulate inventory. This enables the estimation of retail competition in single category demand analyses. We also adapt the storable dynamic discrete choice model to consider goods that are durable, and incorporate an optimal durable good replacement problem. While storability and durability have been considered separately in other research, we are the first to combine these two features. We also adapt the model to the case of tied goods which incorporates the feature that the choice between aftermarket alternatives is conditional on the past purchase of primary goods. These modeling contributions may be applied to other single-category analyses of retail competition, other incidences of durable and storable goods (e.g., light bulbs), other tied goods (e.g., technology products like printers and cartridges; and CPG products like Brita\textsuperscript{©} water filters and Swiffer\textsuperscript{©}), other complementary goods (e.g., IPod\textsuperscript{©} and ITunes\textsuperscript{©} songs), and other industries where aftermarket choice is contingent upon the portfolio of primary goods owned (e.g., computers and software).

From a substantive perspective, the paper advances the literature on tied goods. We are the first to consider the problem of selling tied goods through an oligopolistic retail channel. We note that double marginalization may be particularly problematic due to the

\footnote{Note that our measures of retail competition are lower bounds because we consider inter-channel competition, while competition among stores within the channel could also reduce retail pricing power.}
common desire of manufacturers to maintain low primary good prices. Furthermore, we identify what we call a “cross-product” horizontal externality in which the pricing decisions for one product in one retail channel can actually decrease demand for another product in a different retail channel if the goods are complementary and there is limited retail competition in either channel. The potential for these channel problems are relevant for managers and depend on the level or retail competition. While many marketing researchers have either assumed monopolistic retailers or found limited retail competition, our analysis suggests that retail competition in one of the most commonly noted tied goods industries is, in fact, sufficient to prevent any retailer from reducing adoption of the razor below the level that would occur in a fully integrated 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

Our goal is to develop a model of demand for tied goods sold through competing distribution channels. Our approach is to estimate the model on panel data on consumers’ purchases of razors and blades across various distribution channels, and to then use the model estimates to simulate across store substitution effects. To motivate the model formulation, we first discuss the important considerations that the model is intended to capture. First, purchase of the primary good (i.e., razors) is often temporally separated from purchases of the aftermarket good (blades). Consequently, consumers’ decisions regarding purchase of the primary good are based on expectations about the future prices of the aftermarket goods. We incorporate these expectations formally by deriving the demand system from an underlying discrete-choice model of dynamic consumer decision-making under uncertainty, in which consumers internalize the effect of future prices in deciding the timing and quality of their purchases. The ability of the discrete-choice framework to handle changing choice-sets also enables us to incorporate another important aspect of the tied goods problem, which is that the consumer’s choice set is dependent on their past purchase behavior.

Vertical restraints may be desired by the manufacturer in other tied goods contexts. Two-part tariffs or quantity discounts can solve the double marginalization problem. Resale price maintenance can help curtail the extent of the cross-horizontal externality. Exploring which vertical restraint(s) achieves the efficient outcome for the manufacturer would be an interesting avenue for future research.
Second, competing distribution channels imply that customers form expectations over the prices at each channel and then optimally select a channel at which to purchase. In the context of items that are purchased as part of a much larger basket, the good itself may not drive the channel choice. Rather, consumers may need to wait until their next visit to a low priced channel to take advantage of lower expected prices. Consequently, a channel’s ability to attract a consumer away from a competitor is driven by the frequency with which the consumer visits the channel. We capture this aspect of channel market power by incorporating consumer expectations about the probabilities of future retail visits into the model of demand. This also implies that cross-channel substitution is a form of intertemporal substitution because consumers do not simultaneously visit multiple channels.\footnote{This formulation of intertemporal substitution across and within channels captures the concept of temporal and spatial dimensions of price search as considered by Gauri, Sudhir and Talukdar (2007).} The storability of the products implies that by capturing cross-channel substitution, we also endogenize the store at which consumers choose to accumulate inventory.

**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 \( r = 1, \ldots, R \), 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 \( j \in \{1, \ldots, J_0k\} \) only contains disposable blades. 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, \( K = 3 \).

**Consumers and States**

We first describe the state variables that describe the consumer’s problem. 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 outlet visited and the
different razor types and is currently using disposable blades).

<table>
<thead>
<tr>
<th>Ω</th>
<th>ρ</th>
<th>Razor technologies owned</th>
<th>Blade type currently owned</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,1,1)</td>
<td>0</td>
<td>X</td>
<td>Mach Sensor Schick</td>
</tr>
<tr>
<td>(1,1,1)</td>
<td>1</td>
<td>X</td>
<td>Mach Sensor Schick</td>
</tr>
<tr>
<td>(1,1,1)</td>
<td>0</td>
<td>X</td>
<td>Mach Sensor Schick</td>
</tr>
</tbody>
</table>

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 will 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.\[9\]

Level of depreciation (z):

8 This allows us to rationalize the observed pattern in the data of consumers switching between blade
packs without purchasing razors in between.
9 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 Erdem, Imai
and Keane (2003) 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.
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 $\gamma z$, where $\gamma$ 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.

Store visited ($k$):

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

Prices ($p_k$):

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 $p_k$, where $p_k = \{p_{11k}, \ldots, p_{R,Jrk}\}$.

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

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 $s$ 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 $j^* \in \{1, \ldots, J_{rk}\}$

b. Stay with the current blade type $\rho$ and not buy any blades, or,

c. Stay with the current blade type $\rho$ and buy the blade pack $j^* \in \{1, \ldots, J_{rk}\}$ 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.

Each week, the consumer chooses the actions that maximize the expected present discounted value of current and future payoffs associated with those actions. We model the consumers’ choice as the outcome of a discrete-choice utility maximization problem. We discuss the specific utility functions used below.

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 \( \rho \) to \( r \) by buying a pack of type \( r \) from any of the \( K \) stores 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, \( \gamma \), (since \( z = 1 \)). We then write the current utility from purchase of pack \( j \) of type \( r \neq \rho \) as,

\[
 u_{\eta}(\rho, b, p_k) = \gamma - \alpha p_{\eta j} + \lambda \gamma b + \epsilon \eta
\]

Here, \( \alpha \) is the price sensitivity, \( p_{\eta j} \) is the price of the pack, and \( \epsilon \) is an iid extreme-value shock to utility, unobservable to the econometrician. We include the term \( \lambda \gamma b \) to allow the consumer to receive a salvage value for the stock of blades from the old technology, \( \rho \). 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.\(^{10}\) \( \lambda \) 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.

Current utility of staying with current blade type \( \rho \) and not buying any blades

\(^{10}\) Alternatively, we could exclude this term, thus implicitly assuming that the consumer costlessly disposes the entire inventory of blades associated with the old technology \( \rho \). This assumption has a disadvantage that it imposes a priori large costs of switching razors when owning many blades.
If a consumer stays with the current blade type \( \rho \) and chooses to not buy any blades, he obtains utility:

\[
u_0(\rho, b, z) = \gamma_\rho z + \epsilon_0\]

Thus, if a consumer does not purchase any razors or blades, he obtains the per-period utility of using his current blade as long as \( z > 0 \). 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 \( \rho \) and buying blade pack \( j \)**

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

\[
u_P(\rho, b, z, p_j) = \gamma_\rho z - \alpha \rho \beta_k + \epsilon_{\rho j}, j > 1\]

In contrast to switching blade types, the consumer does not receive a salvage value when buying refills.

Finally, the consumer also faces a decision of whether to replace the blade currently being used. Replacement of the blade does not involve any purchases as long as some blades are available in inventory. The decision to replace a dull blade 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. We present this below.

**State Transitions**

Dynamics arise in this model through the transitions of the state variables \( \{\Omega, \rho, z, b, k, p_k\} \). The key dynamics of interest are generated by the state variables \( \Omega, \rho, z, b \). 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 \( y_k = \{y_{0k}, \ldots, y_{10k}, \ldots, y_{1k}, \ldots, y_{01k}, \ldots\} \). \( y_{00} \) is an indicator variable denoting no purchase (of razors or blades) while at store \( k \); the indicators \( y_{rkj} \) \( (r \neq \rho) \) denote purchase of pack \( j \) for a new technology \( r \); and the indicators \( y_{\rho j} \) denote purchase of blade pack \( j \) of the currently owned blade type \( \rho \). Finally, \( y_0 \) is an indicator of no store visit. Due to the discrete choice aspect of the problem, we have that,
\[ y_0 + \sum_{k=1}^{K} y_{k0} + \sum_{k=1}^{K} \sum_{r=0}^{R} \sum_{j=1}^{J} y_{krj} = 1 \]

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:

\[ \rho_r' = \begin{cases} 
r & \text{if } \sum_{j} y_{krj} > 0, r \neq \rho, k > 0 \\
\rho & \text{otherwise} 
\end{cases} \quad (6) \]

Thus, purchase of any pack, whether containing a razor plus blades, or only blades of a new technology \( r \), switches \( \rho \) 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 \( \delta z \), where \( \delta \) 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 \( \delta \). 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 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 $\delta$ to be different for disposables and non-disposables.

**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:

$$b'(r,j) = q_{rjk} - 1, \text{ if } y_{rjk} = 1, r \neq \rho, k > 0$$

where $q_{rjk}$ is the number of blades included in pack $j$. We subtract one because a consumer immediately starts using one of the new blades when purchased. In the following period, the consumer will therefore have $q_{rjk} - 1$ unused blades and one blade with a level of $\delta$.

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, $q_{\rhojk}$:

$$b'(\rho, j | c) = \begin{cases} b - c + q_{\rhojk} & \text{if } y_{\rhojk} = 1, k > 0, \ (b - c + q_{\rho}) \leq B \\ B & \text{if } y_{\rhojk} = 1, k > 0, \ (b - c + q_{\rho}) > B \end{cases}$$

In (9.1) above, we restrict the number of blades to be less than $B$ to keep the state space for the dynamic problem bounded. This is not a binding constraint because we set $B$ such that it is greater than or equal to the maximum inventory a consumer would hold. The transition of the blade stock if the consumer does not purchase a razor or any blades is analogous to (9.1),

$$b'(\rho, 0 | c) = \begin{cases} b - c & \text{if } \ (y_{k0} = 1 | y_0 = 1), k > 0, \ (b - c) \leq B \\ B & \text{if } \ (y_{k0} = 1 | y_0 = 1), k > 0, \ (b - c) > B \end{cases}$$

**Store visited**

We assume that each consumer has a probability $\phi_k$ 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 $\phi$. Given that razors and blades form a small part of the overall shopping budget, our implicit assumption is that the store choice probabilities are not shifted by prices (we find support for this assumption in the data). Letting $I(k)$ denote an indicator for whether the consumer visits retail outlet $k$, we write the outlet visit probabilities as,
\[ \phi_k' = \Pr(I'(k) = 1) = \frac{\exp(\mu_k + v I(k))}{\sum_{m=1}^{K} \exp(\mu_m + v I(m))} \]  

where, \( v \) and \( \mu_k \) 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 be bi-modally distributed (i.e., a base price and discount price, as frequently observed in retail scanner panel data). There are consequently \( 2^{\sum J} \) 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, \( \psi_{km} \) where \( m \) denotes a given combination of prices (in our empirical application, \( m \) is of the order of 444 at the drug/grocery channel).

Implicitly, prices are allowed to be flexibly contemporaneously correlated across products, but assumed to be independent over time. We make the independence assumption to reduce the size of the state space for the dynamic problem, and to reduce the computational burden of the estimator.

**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). The alternative-specific value functions encapsulate the expected present-discounted utility associated with the choice of each alternative, when the consumer continues to make decisions optimally in the future. The alternative-specific value functions are the dynamic analogues of choice-specific payoffs in static discrete-choice demand systems. 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,
The alternative-specific value function encapsulates the current utility from switching to a new razor today as well as the option value associated with potential actions that could be taken tomorrow, conditional on the decision to switch to a new razor today. The option value includes a) the surplus the consumer would obtain from the stream of consumption associated with the optimal purchase and usage of blades of the new razor in the future, and b) the potential surplus from changing to a different razor in the future, if that turns out to be the better option later. Since future states are uncertain, the option value needs to incorporate the uncertainty associated with the surplus accruing in the future. Further, all payoffs associated with future actions have to be discounted to present utility terms by the discount factor $\beta$.

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 $\phi_0$) and staying on with the new blades $r$ tomorrow. A switch to a new blade type today also encapsulates the discounted value of potentially visiting each outlet $k$ with probability $\phi_k$ 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 $r' \neq r$, via purchase of a pack $j \in (1.., r_j)$ (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, $p_k$, and over the distribution of unobservable shocks to utility that could be realized in the future (i.e., $\epsilon$-s), evaluated at the value of the other state variables tomorrow $s' = \{\Omega', \rho' = r, z' = \delta b' = b'(r,j), k, p_k\}$.

**Value of staying with current blade type $\rho$ and buying blade pack $j$**

The value to a consumer from staying with the current blade type and buying blade pack $j$ implicitly depends on whether or not the consumer changes blades, $c$. 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 after buying the blade pack $j$, he changes the blade currently.
used, the consumer starts off with a new blade next period with \( z' = 1 \), giving him expected future payoff,

\[
EV_{\rho_j}(\rho, z, b, \Omega | c = 1) = \beta \phi_0 E \left[ V_0(\rho, 1, b' (\rho, j | c = 1), \Omega) + \varepsilon_0 | b > 0 \right]
\]

\[
+ \beta \sum_{t=0}^{\infty} \phi \sum_{\psi_{mk}} \text{E} \max \left\{ \left[ W_{\rho_j}(\rho, 1, b' (\rho, j | c = 1), \Omega, p_k) + \varepsilon_{\rho_j} \right]_{j=1}^{l_{\rho_k}}, \left[ V_{\rho_j}(\rho, 1, b' (\rho, j | c = 1), \Omega, p_k) + \varepsilon_{\rho_j} \right]_{j=1}^{l_{\rho_{k'}}} \right\}
\]

(12)

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

\[
EV_{\rho_j}(\rho, z, b, \Omega | c = 0) = \beta \phi_0 E \left[ V_0(\rho, \delta z, b' (\rho, j | c = 0), \Omega) + \varepsilon_0 \right]
\]

\[
+ \beta \sum_{t=0}^{\infty} \phi \sum_{\psi_{mk}} \text{E} \max \left\{ \left[ W_{\rho_j}(\rho, \delta z, b' (\rho, j | c = 0), \Omega, p_k) + \varepsilon_{\rho_j} \right]_{j=1}^{l_{\rho_k}}, \left[ V_{\rho_j}(\rho, \delta z, b' (\rho, j | c = 0), \Omega, p_k) + \varepsilon_{\rho_j} \right]_{j=1}^{l_{\rho_{k'}}} \right\}
\]

(13)

The expected value of the 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:

\[
V_{\rho_j}(\rho, z, b, \Omega, p_k) = (\gamma_{\rho, z} - \alpha_{p_{\rho_k}}) + \max_c \left\{ \frac{EV_{\rho_j}(\rho, z, b, \Omega | c = 1), \left[ EV_{\rho_j}(\rho, z, b, \Omega | c = 0) \right] \right\}
\]

(14)

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.

**Value of staying with current blade type \( \rho \) and not buying any blades**

Finally, the value to a consumer from staying with the current blade type and not buying any blades also depends on the choice of \( c \). To choose \( c \), the consumer compares \( EV_{\rho}(\rho, z, b, \Omega | c = 1) \) to \( EV_{\rho}(\rho, z, b, \Omega | c = 0) \), where the expressions for the payoffs \( \{ EV_{\rho}(\rho, z, b, \Omega | c = 1), \} \) are exactly the same as \( \{ EV_{\rho}(\rho, z, b, \Omega | c = 1), \} \), except that \( b'(\rho, j) \) in (13) and (14) are replaced with \( b'(\rho, 0) \). Also, note that the discounted future value
for changing razors, $EV_{\rho|c=1}$ is conditional on the consumer having remaining unused blades in stock ($b > 0$). If this is not the case the consumer can only receive $EV_{\rho|c=0}$. Given these values, we can now write the value function for staying with the current blade type and not buying any blades as:

$$V_0(\rho, z, b, \Omega, \rho_0) = \begin{cases} \gamma z + \max \left\{ EV_{\rho|c=1}, EV_{\rho|c=0} \right\} & \text{if } b > 0 \\ \gamma z + EV_{\rho|c=0} & \text{otherwise} \end{cases} \quad (15)$$

The system of equations 11-15 above jointly define the value functions that determine the optimal choices for the consumer. Following Rust (1987), the integration with respect to the iid extreme value shocks can be performed analytically. The resulting value functions are defined recursively and cannot however be solved for analytically. Following standard arguments (e.g., Rust 1996), the resulting system of functional equations are a contraction mapping for the corresponding alternative specific value functions. Hence, iterating on the system from an initial guess for the values is guaranteed to converge to a unique fixed point. This facilitates numerical computation. We approximate the value functions numerically over a grid of the state variables, and compute them by value function iteration. To evaluate 14 and 15, we separately evaluate the expected payoffs, and pick the maximum. Given three outlet-types and the set of pack sizes available at each store (see table 2), the system of equations 11-15 above corresponding to our empirical application is of dimension 28 (i.e., we compute 28 alternative-specific value functions). Due to the size of the problem, we use a finite horizon of sixty weeks to approximate the solution. One solution of the values took about 134 seconds on an Intel Xeon architecture PC.

**Discussion**

We have modeled razor-ownership as an absorbing state, i.e., once a consumer has adopted the primary good, he is assumed to have it forever. This assumption seems a priori reasonable for razors. 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. We have also chosen to model the replacement decision of the aftermarket good by modeling the blade-depreciation rate, $z$. 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 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. Finally, we have also chosen to explicitly incorporate disposables, a non-tied technology, into the model. While this aspect is specific to the razor and blades market, the existence of non-tied substitutes is a common feature of many tied good markets (e.g., disposable cameras that come pre-loaded with film). In this sense, our view is that while the framework presented above incorporates aspects specific to the razor and blades market, it is flexible enough to capture interesting demand dynamics in many tied good contexts.

3 Maximum likelihood estimation

We estimate the parameters of the model via maximum likelihood. Computation of the likelihood requires computation of the value functions that define payoffs associated with each action. The estimation procedure incorporates a value function iteration procedure that numerically evaluates the above value functions for each guess of the parameter vector. 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, $\theta_i = \{y^r, \delta_{disp}, \delta_{nondisp}, \alpha_i, \lambda_i, (\phi^k)_{t=1}, (\upsilon)\}$ (here, we allow the depreciation rate $\delta$ to vary between disposables and non-disposables). The discount factor, $\beta$, is not estimated, and is set to 0.998.

Heterogeneity

The econometric model generated by the tied good problem above implicitly generates state dependence in demand due to the dependence over time in the unobserved blade inventories of consumers. As in any state-dependent demand model, specifying a rich heterogeneity distribution is critical for properly sorting this state-dependence from time-invariant consumer-specific unobserved factors that generate persistence over time. We allow for a multidimensional continuous heterogeneity distribution using random coefficients. We specify
the parameters, $\theta_i$, to be multivariate normally distributed across consumers with a covariance matrix defined by $\Sigma$:

$$\theta_i = \bar{\theta} + \Gamma \eta_i$$  \hspace{1cm} (16)

where $\Gamma$ is the Cholesky decomposition of the covariance matrix, $\Sigma$, and $\eta$ is a vector in which each element is distributed standard normal. An appealing aspect of the specification above is that it accommodates correlation in the outlet visit propensities with consumer price sensitivities and usage utilities. In addition to providing considerable model flexibility, this enables us to endogenize the taste distribution of consumers who visit and self-select into stockpiling blades at each retail outlet. 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. The model implies that the probabilities are:

$$\Pr(y_{i,tkt} = 1) = \frac{\phi_{it} e^{V_i(\rho_{it}, z_{it}, b_{it}, k_{it}, \Omega_{it}, p_{it})}}{\Lambda_{it}}$$

$$\Pr(y_{i,\rho_{it},jk_t} = 1) = \frac{\phi_{it} e^{V_i(\rho_{it}, z_{it}, b_{it}, k_{it}, \Omega_{it}, p_{it})}}{\Lambda_{it}}$$

$$\Pr(y_{ijkt} = 1) = \frac{\phi_{it} e^{V_i(\rho_{it}, z_{it}, b_{it}, k_{it}, \Omega_{it}, p_{it})}}{\Lambda_{it}}$$

$$\Pr(y_{i,t} = 1) = \phi_{it}$$

where,

$$\Lambda_{it} = e^{V_i(\rho_{it}, z_{it}, b_{it}, k_{it}, \Omega_{it}, p_{it})} + \sum_{j=1}^{l_{jt}} e^{V_i(\rho_{it}, z_{it}, b_{it}, k_{it}, \Omega_{it}, p_{it})} + \sum_{r \neq t} \sum_{j=1}^{l_{jt}} e^{V_r(\rho_{it}, z_{it}, b_{it}, k_{it}, \Omega_{it}, p_{it})}$$

and, $\rho_{it}, z_{it}, b_{it}, k_{it}, \Omega_{it}, p_{it}$ are state variables for the individual each week. The likelihood function for a given individual in the data for a week $t$ is therefore:

$$\ell_t \left( y_{i,t}, y_{i,1t}, \ldots, y_{i,kt} \mid \rho_{it}, z_{it}, b_{it}, k_{it}, \Omega_{it}, p_{it}, \theta \right)$$

$$= \left[ \Pr(y_{i,t} = 1) \right]^{y_{i,t}} \prod_{k=1}^{K} \left[ \Pr(y_{i,0kt} = 1) \right]^{y_{i,0kt}} \prod_{j=1}^{l_{jt}} \left[ \Pr(y_{i,\rho_{it},jk_t} = 1) \right]^{y_{i,\rho_{it},jk_t}} \prod_{r \neq t} \prod_{j=1}^{l_{jt}} \Pr(y_{ijkt} = 1)$$

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 $\delta$. These are specified in table 3, where we present the estimation results.
The likelihood of the data for the individual across all weeks is then,

\[
L_i(\Theta) = \int \prod_{t=1}^{T} \ell_t(y_{0t}, y_{1t}, \ldots, y_{kt} \mid \rho_i, z_i, b_i, k_i, \Omega_i, p_{kt}; \theta) \pi(\rho_{0i}, z_{0i}, b_{0i}, \Omega_{0i})] d\Phi(\theta_i; \Theta) (19)
\]

where \(\rho_{0i}, z_{0i}, b_{0i}, \Omega_{0i}\) are the initial current razor, current blade depreciation level, blade stock, and razor ownership, and \(\pi(\rho_{0i}, z_{0i}, b_{0i}, \Omega_{0i})\) 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, \(q_{rit}\), and the initial states.\(^{12}\)

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.

Specifically, to obtain the initial razor ownership \(\Omega_{0i}\), 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 \(\rho_{0i}\) as follows. If the first observed purchase for a consumer is a blade pack, we set \(\rho_{0i}\) 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 \(\delta\) is the current guess of the depreciation rate.

\(^{12}\)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.
for the consumer.\footnote{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 $(1-\delta)$.} Finally, we assume consumers start with a new blade at the beginning of the data – i.e., we set $z_{0i} = 1$.

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 Ackerberg (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 (see Ackerberg 2001).

**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, $\gamma$, 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, $\lambda$, are identified from the time-elapsed 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, we infer that $\lambda$ 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 heterogenous 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 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.
4 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.

The period of our data includes the Gillette Mach razor, the Gillette Sensor razor, as well as the Schick Xtreme3 and Tracer (it does not include the Schick Quattro which was introduced in late 2003). Among these, the Gillette Mach razor was the dominant razor technology. While Gillette also sold the Sensor razor model during this time period, we see primarily refill purchases of this technology to established customers (195 refill cartridges were sold, while only 5 razors were sold across consumers in a 1 year period). Further, the share of Schick Xtreme3 and Tracer razors was very limited.\textsuperscript{14} Given that Gillette Mach is the only dominant razor and blade system in the market at this time, we restrict the consideration of tied good dynamics to the Gillette Mach technology. This implies that when modeling consumer choices, Mach refill packs only enter the consideration set after a Mach razor is purchased to reflect the tied good dynamics, but we allow consumers to choose from all Schick and Sensor packs regardless of their past purchase history.\textsuperscript{15}

Data cleaning

In order to obtain a dataset suitable for estimation, we undertook several steps to clean the raw data. First, we restricted attention to households that purchased at least one male razor or blade pack during the period of the data. Second, our desire to model discrete choices over packs required us to define a finite set of possible pack sizes and to group each observed

\textsuperscript{14} Our data contain only 12 purchases of the Xtreme3 razor, 3 purchases of the Tracer and 2 additional purchases of other Schick razors. These 17 Schick razor purchases compare to 71 purchases of the Gillette Mach razor during our data. Furthermore Schick only sold 69 cartridge packs during this time period compared to 438 cartridge packs sold for Gillette Mach.

\textsuperscript{15} Allowing customers to buy Schick and Sensor refill packs regardless of observed history facilitates our estimation in two ways. First, modeling tied good dynamics of two additional technologies implies an increase in the computational time of our model that is much more costly than the benefit of accounting for the limited razor sales for these technologies. Second, since most Sensor and Schick razors were likely purchased before our data begin, modeling their tied good dynamics would rely heavily on assumptions about initial conditions because there are very few observations to inform us about the incentives to purchase these razors.
purchase within this set of mutually exclusive choices. This task was easiest within the Gillette technologies because we were able to define a choice that corresponded to each pack size. Furthermore, large pack sizes are typically twice the size of smaller pack sizes (Machs are sold in 4 and 8 packs while Sensors are sold in 5 and 10 packs) such that a purchase of two small packs could easily be coded as a purchase of a single large pack. The Schick purchases were grouped into 4 and 8 packs based on the closest corresponding number of blades purchased. Disposable purchases were grouped into 4 packs, 8 packs, 20 packs and 60 packs based on which of these pack sizes most closely approximated the total number of blades purchased during the week. While one might prefer to accommodate this large number of pack sizes with a continuous choice of the number of blades, the reality is that blades are sold in discrete quantities and modeling choice over discrete pack sizes is important to accommodate the nonlinear pricing implied by per-blade prices that decrease with the size of the pack. Furthermore, the dynamic discrete choice literature is not well-suited to modeling choice over continuous variables. Consequently, other researchers who have considered stockpiling behavior (Erdem, Imai and Keane, 2003 and Hendel and Nevo, 2006) have also modeled discrete choices over a finite number of pack sizes.

Third, we also model consumers to make a discrete choice over the channel visited. When we observe a consumer visiting multiple channels in the same week, we assign them to a single channel based on the following rule. If we observed them purchasing a razor or blade pack in a given week, we assume they only visited the channel in which they made that purchase. If we do not observe them purchasing in the category, we assign them to the channels in the following order: Club Stores, Mass Merchandiser; Drug/Grocery. For example, if a customer visits a grocery store and a club store and does not purchase in the category, we assume they visited a club store. The reason for this is that club stores and mass merchandisers are visited rarely, while grocery stores are visited quite often. We would rather

---

16 Six 10-pack cartridges were grouped into the 8-pack choice, nine 5-pack cartridges were grouped into the 4-pack choice, ten Schick razors which typically contained 3 blades were grouped into the four pack choice, two Schick razors which contained 4 blades were also grouped into the 4-pack choice, and the remaining five Schick razors also involved a 4-pack refill purchase and were consequently classified as 8-pack purchases.

17 4-packs included as many as six disposable blades. 8-packs included as many as twelve disposable blades and also included packs with as few as 4 blades if multiple blade packs were purchased during the week. 20-packs typically include packs with 24 or fewer blades, though there was one 40 pack included in this group. 60-packs included a blade pack as large as 74. In our model, this measurement error in the number of blades actually purchased will be absorbed in the same manner that random variation in blade usage is accommodated.
slightly over represent the small visit share of club stores and mass merchandisers than over-
represent the already large share of visits to grocery and drug stores.\textsuperscript{18}

Finally, we face a data issue in that, for each consumer, we observe prices only for the
product actually purchased. Since several club stores and mass merchandisers do not share
price and quantity data, the store-level shelf prices at these outlets are not available, and
cannot be procured even by syndicated data vendors. Hence, from the perspective of a
researcher using these data, the prices of competing products available at stores need to be
inferred. A potential methodology to infer the prices of the other products available in a retail
outlet each week is to infer these from the observed prices paid by other consumers. We found
that the data are too thin at the product-outlet-week level to make this robust and viable.
Instead, for each outlet, 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. Table 1 indicates that the consumers in this panel are
most likely to visit a drug or grocery store each week (44\% of visits) followed by mass
merchandisers (30\%) and club stores (11\%). There is a 15\% chance that no store is visited each
week.

\textbf{--- Tables 1 and Figure 2 here ---}

The bottom panel of Table 1 presents the distribution of razor ownership across
consumers and weeks. Disposables have the highest ownership share, mainly reflecting the
purchases of large pack sizes of disposables that we see in our data. Among non-disposables,
Gillette-Mach, the focal razor of our analysis, has the highest ownership share, which reflects
the well known strong market position that Gillette holds in the razor and blade market.
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

\textsuperscript{18} An alternative is to move to a model specified at the daily or bi-weekly level, which would be
unnecessarily burdensome.
razor ownership by consumers. The share of observations with other blade types in inventory is spread across Sensor, Schick and disposables having 9, 4, and 58% shares respectively.

To get a sense of the extent of heterogeneity in usage across the consumers in the data, we plot in Figure 2 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. From figure 2, we see that while some customers only purchased 3 blades across the 56 weeks (perhaps, preferring to consume the outside option, an electric shaver, most of the time), there are many that have high usage needs, including 45 consumers that purchased more than 60 blades across the 56 weeks of the data.

**Stocking and Sales of Pack Sizes by Channel**

In Table 2a, we list the pack sizes available at each retail outlet, as well as the number of purchases of each of the pack sizes 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. As mentioned previously, our specification of heterogeneity allows for correlation of preference parameters with store visit propensities, and is able to capture this aspect of retail substitution. 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.

--- Tables 2a and 2b here ---

It is useful to analyze the column totals in Table 2a. About half of all purchases (796 of 1,565) are made at grocery or drug stores. 42% of sales are at mass merchandisers and only seven percent are at club stores. Interestingly, club stores form a significant portion of the blade market, making up over 18% of all blade sales, despite only seven percent of all purchases. Looking across the columns in Table 2, we also see that the total sales by technology resemble the ordering of the razors held as depicted in Table 1. Disposables are the most commonly purchased product, with 55% of all purchases. This partly reflects the large pack sizes and the low per-unit prices of disposables. Mach sales are 31% and Sensor 10, while Schick is at about 4%.
Pack Size Prices Across Channels

Table 2b lists the prices of each of the pack sizes at each channel from Table 2a. 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. For instance, a Mach 3 pack that contains a razor plus 3 blades costs about $7.72 on average at a grocery store, compared to $7.77 for a Mach 4 pack containing 4 blades and no razor. Table 2b 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. For example, looking at Gillette-Sensor, we see that the per-blade prices are lower for 10 packs ($1.30 at grocery outlets, and $1.17 at mass merchandisers on average) versus 5 packs ($1.37 at grocery outlets and $1.23 at mass merchandisers on average). This aspect of price discrimination is accommodated in our empirical model since we explicitly model choices over pack sizes.

Descriptive Statistics about Channel Switching

A crucial piece of our analysis involves the measurement of retail competition. 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. MK stores have a greater mass of consumers at 0% visit probabilities, but most of the mass has a greater than 10% visit probability with a thick right tail. Club Stores have a significant number of consumers with 0% visit probabilities, but the mode is still greater than 0%. There is also substantial mass at 10% and 20%, though the right tail is thin.

While Figure 3 reveals that customers have opportunities to buy razors and blades at different store types, we now consider whether they actually do buy across different store types. 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. The vertical axis measures the density of consumers that purchased a given fraction of
blades at the channel. We see that there is a significant mass of consumers that buy either all or none of their blades at DG and MK stores. There is also a positive mass of consumers buying blades in these channels with some probability between 0 and 1. At W 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 clubs stores with a proportion greater than zero and less than 40%. Together, Figures 3 and 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.

5 Model Estimates

We present the estimates of our dynamic random coefficients model defined above. Models without heterogeneity, or with myopic consumers, are unrealistic in this context. Ignoring heterogeneity leads to the common positive bias in state-dependence. This is problematic because positive state dependence implies that consumers would be more likely to purchase the greater their inventory. A myopic consumer version of this demand model would perform particularly poor, since it implies that prices of blades do not affect razor demand and that the additional utility from say, an 8 pack of blades, relative to a 4 pack must be obtained in the purchasing week, despite the fact that it would be difficult to use all 8 blades in 7 days. The estimates from the random coefficients model are reported in Table 3. The brand utilities from Table 3 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 respectively, calculated by taking the exponents of the reported values and dividing by one plus this exponent, are 0.63 and 0.002. Thus after a single week of usage, a disposable is almost fully depreciated, while Mach and other branded razors retain over 60% of their value after the first week. A take away in terms of comparing products from the estimates is that disposable blades depreciate much faster than non-disposable blades, which is consistent with the conventional wisdom.

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 16% (transforming -1.633 by its exponent divided by one plus its exponent). This implies that while switching blade types, an individual obtains 16% of the un-depreciated
per-period utility associated with the current blade type owned. While individuals may not actually sell their blades, we can think of this value as capturing the discounted present value of future utility that consumers would have obtained by instead storing these blades, and using them at some distant future date.

The price coefficient is 3.642 (i.e., exp(1.293)) 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 outlet visit parameters represent the average probability of visiting each of store types 1-3 (grocery, drug and wholesales/club stores respectively). 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 outlet 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 2b, 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 additional important determinant of the market power of the retail channels. In particular, the store choice 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 competition, while negative correlations indicate little inter-channel competition. These correlations are presented in the second panel of Table 4. We see that mass merchandising channels are substitutable with both drug/grocery and club
stores, based on positive correlations in preferences of 0.60 and 0.38 respectively. Yet, drug/grocery and club stores are not very substitutable with one another, indicated by a correlation of 0.14. 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 less price sensitive consumers.

Table 4 also reports the estimated correlations in usage utilities, as well as the correlations of usage utilities with the depreciation and the price sensitivity. We see that there is a strong positive correlation between the usage utilities for Gillette razors (i.e., Gillette-Mach and Gillette-Sensor). Interestingly, Sensor (the older Gillette technology) is positively correlated with utility for disposables, while other branded razors are not. We also see that customers with strong preferences for Mach razors tend to be less price sensitive, but at the same time, also tend to have lower depreciation rates for their blades (the positive correlation implies that $\gamma_{MACH}$ is positively correlated with $\delta$; recall, the higher the $\delta$, the lower is the depreciation).

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 5, 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 5, we see that the model implies that the more unused blades a consumer has, the less likely he is to buy any blades. We also see that the value of purchasing more blades is decreasing with the blade stock at a faster rate for Mach blade packs (i.e., the lines are steeper for Mach packs compared to non-Mach packs). Further, within Mach blade packs, the value of purchasing more blades is decreasing with the blade stock at a faster rate for the larger Mach blade packs (for instance, the lines are steeper for a Mach 8 pack relative to a Mach 4 pack at a drug/grocery store, or for a Mach 20 pack relative to a Mach 16 pack at a club store). Both effects are intuitive: since this is a customer who already owns Mach blades, the value of additional Mach blades is lower the more blades he
owns in inventory. Further, this decrease is higher when there are more blades in the pack to be purchased.

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 $1 disincentive to purchase either of the Mach blade packs at a drug/grocery store. 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 $0.43, 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 $5 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.

Finally, the same consumer may prefer to buy razors at a different channel than he prefers to buy blades. Although not reported, it is easy to compute the implied surplus he would obtain if he were to buy a Mach razor pack at a drug/grocery store, when his state is such that he does not own a Mach razor. This surplus is about $4. The model thus implies that this consumer has a strong incentive to buy Mach razors at a drug/grocery store, and to subsequently buy large blade packs at club stores. The implied differences in incentives to purchase products across channels illustrates the sense in which interchannel and intertemporal substitution is captured by our proposed model.

**Discussion**

To summarize the demand estimates, we find strong evidence for heterogeneity in the data, as well as strong 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.

6 Measuring Retail Competition

We now return to the research question posed at the beginning of the paper, i.e., whether or not there is sufficient retail competition to allow manufacturers of tied goods to maintain desired relative pricing between primary (razors) and aftermarket goods (blades). While double marginalization can be harmful to manufacturers generally, it is particularly problematic in the case of tied goods if it leads to reduced adoption of the primary good, which is the acquisition device used to obtain the stream of future revenue from aftermarket sales. In other words, we seek to quantify whether a retailer’s increase in razor prices reduces aggregate razor demand and, if so, how much the reduced razor adoption reduces demand for the complementary blades.

Note that our primary interest here is to measure the long-run effects of price changes. This is important because, while a temporary increase in the razor price may reduce adoption in the short-run, it is likely that the consumer will just postpone adoption of the razor. This may have a minimal impact on subsequent blades sales. In contrast, a permanent price increase that prevents a consumer from ever adopting a razor also eliminates all blade sales that otherwise would have been made to the consumer. We measure these long-run effects by simulating consumer demand in response to a permanent change in the price distribution. Thus, the measures we report incorporate both the demand effect due to the current prices, as well as the substitution across stores and time due to changes in the price expectations of consumers. We first provide a brief description of the procedure to simulate these values using the dynamic structural model we develop.

Simulating Approach for Long-Run Elasticities

We first make $N = 725$ draws from the distribution of estimated consumer preferences, to obtain 725 consumers representative of those observed in the data. For each consumer we simulate 4,000 different purchase histories to simulate the integrals over the extreme-value distribution representing idiosyncratic alternative-and-time-specific unobservables. We begin in week 1 in which the customer has a disposable technology and no blades. We then draw a store choice based on the individual’s estimated preferences and a vector of prices from the joint distribution of prices at that store. Then we take 4,000 draws of the vector of extreme
Combining these draws with the simulated consumers’ value functions found by solving the dynamic programming problem, we simulate 4,000 choices in the first period for each of the 725 consumer types considered. These choices, together with the initial state imply 725X4000 = 2,900,000 different states for the second period. In the second period, we draw a new store choice, a new price vector, and a new set of extreme value draws. We use these to determine 2,900,000 second period choices, which together with the states imply states for the third period. We repeat this for each of 52 weeks, such that we obtain 2,900,000 purchase histories of a one year length. This allows us to determine the expected number of razors and blades purchased at each store.

We then perform the same simulation using the same draws for a case in which the value functions are derived from a solution of the dynamic programming problem in which grocery/drug prices for the Mach razor are 10% greater than the true distribution (i.e., this is our hypothetical price change). This requires the re-computation of the value function for all the consumers to generate a new set of 2,900,000 purchase histories of one year. This second simulation gives us the expected number of razors and blades purchased under the counterfactual of higher Mach razor prices at the grocery and drug channel. Consumer expectations and choice behavior under this scenario thus reflect the permanent increase in the price level of the Mach razor-pack. Our estimates of long-run elasticities are then computed as the percentage difference in the total sales of razors and blades at each outlet, between the two simulations. We repeat this exercise for two other hypothetical price changes (i.e., a change in blade prices at drug/grocery and at club stores).

In the remainder of this section, we use these simulations to explore three aspects of retail substitution. First, we measure the long-run own and cross-channel elasticities of each retail outlet’s razor sales with respect to razor prices. A finding of large own- and cross-effects would imply highly elastic demand, along with high substitutability across stores, suggesting low retail market-power in razors. This would in turn suggest that the extent of the double marginalization problem in razors would be low for the manufacturer. Second, we measure the long-run own and cross-channel elasticities of each retail outlet’s blade sales with respect to blade prices to evaluate retail market power in blades. Finally, we measure the long-run extent to which razor prices at one store affect blade sales at a different store, and vice versa.

---

19 Computation of long-run effects in dynamic models such as ours is numerically intensive. In our simulation above, computation of an elasticity with respect to each price change requires the simulation of 5.8 Million (2X2.9 Million) purchase histories. The computational effort per simulation is linked to the fact that each requires the solution of 1450 dynamic problems (725 consumer types X 2 runs) which each takes 134 seconds, plus simulation of the history conditional on these solutions, which takes an additional 150 seconds per DPP solution. Hence, it takes roughly 57 computing hours to calculate elasticities for one hypothetical price change.
(i.e., the “cross-product” horizontal externalities). Small cross-effects from this simulation would imply that the horizontal externality in tied goods is low. We conclude by discussing the implications of these results for the manufacturers in this market.

**Retail Market Power in Razors**

We report the long-run elasticity estimates (reported as % change in sales due to a 1% increase in price) computed for each of the three hypothetical price changes in Table 5. The first panel reports demand elasticities with respect to a change in the Mach razor price at drug and grocery outlets. We see that the own-price elasticity of demand is -1.215. Cross-price elasticities with respect to the other store types, mass merchandisers and club stores, are 0.921 and 0.768 respectively indicating substitutability across channels. To check whether these substitution effects are “large enough” such that drug/grocery market power is small, we now analyze the net effect of this price change on the total demand for Mach razors from the manufacturer’s point of view. This is presented in the last column of Table 5. We see that a 1% increase in the razor price at the drug grocery channel only reduces manufacturer razor sales by 0.01%. This suggests that the inter-channel competition in the sales of razors is large enough to prevent a drug/grocery channel from substantially reducing the aggregate sales of razors. We do not compute this measure for each of the other two channels due to computational burden. However, the market-power of the other two channels would be less, since the fact that drug and grocery stores sell the most razors and have the highest prices suggests that the strongest retail power for razors would be at this channel.

--- Table 5 here ---

**Retail Market Power in Blades**

We also evaluate two hypothetical blade price increases to evaluate market power in blade sales. The second panel in Table 5 reports elasticities with respect to an increase in the blade price at drug and grocery outlets and the third panel reports elasticities with respect to an increase in the blade price at club stores. We see that own-price elasticities for blades at both channels are strongly negative. Blade demand at club stores is seen to be more elastic in the long-run (-6.800 at a club store compared to -4.136 at a drug/grocery store). The estimated cross-price elasticities also indicate substitution across stores. Looking at the net effect for the manufacturer, it is noteworthy that channels appear to have some ability to reduce the aggregate demand of blades (i.e., a 1% price increase of blades at club stores would reduce manufacturer blade demand by 0.465%). This indicates that club stores may exercise some double marginalization in blades.
Cross-Product Horizontal Externalities

Finally, we examine whether razor prices at one channel significantly shift blade sales at another channel. From Table 5, panel one, we see that an increase in Mach razor price at the drug/grocery channel has only small effects on blade demand at the other retail channels (a 1% price increase generates 0.038% decrease at a mass merchandiser, and a 0.033% decrease at a club store). To see if this has a significant effect on the manufacturer, we estimate the effect on overall manufacturer’s demand for blades. We see that the one percent razor price increase at the drug/grocery store only reduces the aggregate sales of blades by 0.03%. While this effect is rather small, we interestingly see that the reduced blade sales are split quite evenly across all three retail channels. In other words, razor price increases at the drug/grocery channel do decrease blade sales at all channels, but by a rather small amount. This suggests that if the drug/grocery channel had substantial market power in razor sales, it might not only harm the manufacturer through reduced razor and blade sales, but also harm its competitors because they sell a reasonable number of blades to consumers that buy razors at the drug/grocery channel. However, the extent of retail competition in this market prevents this type of “horizontal externality.”

The first row in the last two panels of Table 5 allow us to see if the retail market power in blades generates a cross-product horizontal externality in razors. While the club store is shown in the bottom right to be able to reduce manufacturer demand for blades by raising its prices, it does not generate a large horizontal externality on razor sales (i.e., razors sales only decrease by 0.017% and 0.002% respectively at drug/grocery and mass merchandising channels when the club store raises its prices by one percent). This is likely due to the fact that club stores sell to high volume Mach consumers that may reduce their blade consumption but will not switch technologies due to the price increase. The drug/grocery channel does not appear to generate substantial cross-product horizontal externalities either as evidenced by cross-product cross-channel elasticities of -0.061 and 0.108.  

However, the drug grocery channel itself does internalize some lost razor sales when it increases blade prices due to a slightly larger cross price elasticity of -0.173, which reflects the complementarity of razors and blades sold within the drug/grocery channel.

---

20 One interesting aspect to note is that blade price increases at either channel increase the demand for razor packs at club stores. This is because a club store razor pack contains 7 blades. Therefore, while the razor is a complement to the blades sold elsewhere, the pack itself may be a net complement or substitute because the blades in the pack are substitutes. For example, if a consumer observes a general increase in price for refill packs, it might buy more blades in its initial razor purchase if the razor pack prices are unchanged.
These long-run effects suggest that the downstream retail channel for this industry, while oligopolistic, generally faces highly elastic demand conditions, with significant substitutability across channels, such that retail market-power is low. Overall our results suggest that retail competition in this market is such that manufacturers may be able to pass-through much of their desired pricing through the downstream channel without the need for additional vertical restraints.

7 Conclusions

This paper makes three contributions. First, we identify the horizontal externality aspect of retail competition in tied goods, which has hitherto not been articulated in the literature. Second, we develop a dynamic structural model of tied good demand to measure retail substitution effects, and to calibrate the consequences of downstream competition to the manufacturer. Third, we empirically measure retail competition in a real-world market context, and evaluate from the data whether demand conditions allow manufacturers in this industry to pass low prices on primary goods and desired prices on aftermarket goods through an oligopoly retail channel.

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 competition 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 stores, as well as blades sold across stores. 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.

Finally, the estimates of this structural model for the razor and blades market validated that the competitive structure of the oligopoly retail channel facilitates the level of razor adoption desired by the manufacturer. Without such conditions, the manufacturer may be forced to rely on vertical restraints that have a varied reputation in antitrust due to the potential for inducing anti-competitive behavior.

References


Figure 1
Ratio of Mach Blade to Razor Sales By Channel

Figure 2
Distribution of Blade Usage
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.
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 5. 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

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

Table 2a
Pack Size Purchases by Razor Technology and Store Type

<table>
<thead>
<tr>
<th>Drug and Grocery Stores</th>
<th>Mass Merchandisers</th>
<th>Club Stores</th>
<th>Total</th>
<th>Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gillette - Mach</td>
<td>179</td>
<td>234</td>
<td>73</td>
<td>486</td>
</tr>
<tr>
<td>3pk Razor</td>
<td>26</td>
<td>18</td>
<td>0</td>
<td>44</td>
</tr>
<tr>
<td>7pk Razor</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>4pk Blades</td>
<td>70</td>
<td>123</td>
<td>0</td>
<td>193</td>
</tr>
<tr>
<td>8pk Blades</td>
<td>83</td>
<td>93</td>
<td>0</td>
<td>176</td>
</tr>
<tr>
<td>16pk Blades</td>
<td>0</td>
<td>0</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>20pk Blades</td>
<td>0</td>
<td>0</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>Gillette - Sensor</td>
<td>76</td>
<td>60</td>
<td>17</td>
<td>153</td>
</tr>
<tr>
<td>5pk Blades</td>
<td>47</td>
<td>32</td>
<td>0</td>
<td>79</td>
</tr>
<tr>
<td>10pk Blades</td>
<td>29</td>
<td>28</td>
<td>0</td>
<td>57</td>
</tr>
<tr>
<td>25pk Blades</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Schick</td>
<td>28</td>
<td>40</td>
<td>0</td>
<td>68</td>
</tr>
<tr>
<td>4pk Blades</td>
<td>21</td>
<td>25</td>
<td>0</td>
<td>46</td>
</tr>
<tr>
<td>8pk Blades</td>
<td>7</td>
<td>15</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>Disposables</td>
<td>513</td>
<td>330</td>
<td>15</td>
<td>858</td>
</tr>
<tr>
<td>4pk Blades</td>
<td>184</td>
<td>69</td>
<td>0</td>
<td>253</td>
</tr>
<tr>
<td>8pk Blades</td>
<td>207</td>
<td>156</td>
<td>0</td>
<td>363</td>
</tr>
<tr>
<td>20pk Blades</td>
<td>122</td>
<td>105</td>
<td>0</td>
<td>227</td>
</tr>
<tr>
<td>60pk Blades</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

Total | 796 | 664 | 105 | 1,565 |
Share | 50.9% | 42.4% | 6.7% |
Table 2b
Pack Size Average Prices by Razor Technology and Store Type

<table>
<thead>
<tr>
<th></th>
<th>Drug and Grocery Stores</th>
<th>Mass Merchandisers</th>
<th>Club Stores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gillette - Mach</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3pk Razor</td>
<td>$7.72</td>
<td>$6.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td>($2.57)</td>
<td>($2.24)</td>
<td>$14.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>($2.10)</td>
</tr>
<tr>
<td>7pk Razor</td>
<td>$7.77</td>
<td>$7.41</td>
<td></td>
</tr>
<tr>
<td></td>
<td>($1.94)</td>
<td>($1.85)</td>
<td>$26.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>($1.65)</td>
</tr>
<tr>
<td>4pk Blades</td>
<td>$14.06</td>
<td>$13.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td>($1.76)</td>
<td>($1.67)</td>
<td>$28.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>($1.40)</td>
</tr>
<tr>
<td>16pk Blades</td>
<td>$6.86</td>
<td>$6.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>($1.37)</td>
<td>($1.23)</td>
<td>$25.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>($1.02)</td>
</tr>
<tr>
<td>20pk Blades</td>
<td>$5.00</td>
<td>$4.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>($1.25)</td>
<td>($1.02)</td>
<td></td>
</tr>
<tr>
<td>8pk Blades</td>
<td>$9.99</td>
<td>$8.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>($1.25)</td>
<td>($1.02)</td>
<td></td>
</tr>
<tr>
<td>Disposables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4pk Blades</td>
<td>$0.80</td>
<td>$0.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>($0.20)</td>
<td>($0.23)</td>
<td>$16.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>($0.27)</td>
</tr>
</tbody>
</table>

Notes: Average price per unit in parenthesis
Table 3
Model Estimates

<table>
<thead>
<tr>
<th>Random Coefficients</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Brand Utilities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensor</td>
<td>-2.265</td>
<td>1.150 a</td>
</tr>
<tr>
<td></td>
<td>(0.069)</td>
<td>(0.049)</td>
</tr>
<tr>
<td>Mach</td>
<td>0.266</td>
<td>1.260 a</td>
</tr>
<tr>
<td></td>
<td>(0.078)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>Schick</td>
<td>-3.577</td>
<td>1.212 a</td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>Disposables</td>
<td>-1.197</td>
<td>2.316 a</td>
</tr>
<tr>
<td></td>
<td>(0.146)</td>
<td>(0.125)</td>
</tr>
<tr>
<td><strong>Price Coefficient</strong></td>
<td>1.293</td>
<td>1.386 a</td>
</tr>
<tr>
<td></td>
<td>(0.090)</td>
<td>(0.075)</td>
</tr>
<tr>
<td><strong>Blade Depreciation Rate</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Disposables</td>
<td>0.525</td>
<td>1.334 b</td>
</tr>
<tr>
<td></td>
<td>(0.090)</td>
<td>(0.066)</td>
</tr>
<tr>
<td>Disposables</td>
<td>-6.284</td>
<td>5.462 b</td>
</tr>
<tr>
<td></td>
<td>(0.338)</td>
<td>(0.225)</td>
</tr>
<tr>
<td><strong>Resale Percentage</strong></td>
<td>-1.633</td>
<td>0.702 b</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.034)</td>
</tr>
<tr>
<td><strong>Store 1</strong></td>
<td>1.174</td>
<td>1.140</td>
</tr>
<tr>
<td></td>
<td>(0.068)</td>
<td>(0.050)</td>
</tr>
<tr>
<td><strong>Store 2</strong></td>
<td>0.801</td>
<td>1.622</td>
</tr>
<tr>
<td></td>
<td>(0.093)</td>
<td>(0.079)</td>
</tr>
<tr>
<td><strong>Store 3</strong></td>
<td>-1.592</td>
<td>3.174</td>
</tr>
<tr>
<td></td>
<td>(0.197)</td>
<td>(0.185)</td>
</tr>
<tr>
<td><strong>Store Choice</strong></td>
<td>0.080</td>
<td>0.203</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.009)</td>
</tr>
</tbody>
</table>

Notes: 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 the Heterogeneity Distribution

<table>
<thead>
<tr>
<th>Blade Utilities:</th>
<th>Sensor</th>
<th>Mach</th>
<th>Schick</th>
<th>Disposable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor</td>
<td>1</td>
<td>0.23</td>
<td>-0.54</td>
<td>0.70</td>
</tr>
<tr>
<td>Mach</td>
<td>0.23</td>
<td>1</td>
<td>0.26</td>
<td>-0.07</td>
</tr>
<tr>
<td>Schick</td>
<td>-0.54</td>
<td>0.26</td>
<td>1</td>
<td>-0.60</td>
</tr>
<tr>
<td>Disposable</td>
<td>0.70</td>
<td>-0.07</td>
<td>-0.60</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Branded Depreciation</td>
<td>-0.17</td>
<td>0.60</td>
<td>0.38</td>
<td>-0.440</td>
</tr>
<tr>
<td>Disposable Depreciation</td>
<td>-0.73</td>
<td>-0.26</td>
<td>0.80</td>
<td>-0.55</td>
</tr>
<tr>
<td>Price Sensitivity</td>
<td>0.70</td>
<td>-0.07</td>
<td>-0.63</td>
<td>0.95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Store Choice:</th>
<th>Grocery/Drug</th>
<th>Mass Merch.</th>
<th>Club Stores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grocery/Drug</td>
<td>1</td>
<td>0.60</td>
<td>0.14</td>
</tr>
<tr>
<td>Mass Merch.</td>
<td>0.60</td>
<td>1</td>
<td>0.38</td>
</tr>
<tr>
<td>Club Stores</td>
<td>0.14</td>
<td>0.38</td>
<td>1</td>
</tr>
</tbody>
</table>

Price Sensitivity | 0.12 | 0.05 | -0.32 |
Table 5
Long-Run Demand Elasticities
with Respect to Mach Razor and Blade Prices

<table>
<thead>
<tr>
<th>Retail Channels</th>
<th>Drug/Grocery</th>
<th>Mass Merchandiser</th>
<th>Club Stores</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Drug/Grocery: Mach Razor Pack Price</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mach Razor Pack Demand</td>
<td>-1.215</td>
<td>0.921</td>
<td>0.768</td>
<td>-0.011</td>
</tr>
<tr>
<td>Mach Blade Refill Demand</td>
<td>-0.028</td>
<td>-0.038</td>
<td>-0.033</td>
<td>-0.033</td>
</tr>
<tr>
<td><strong>Drug/Grocery: Mach Blade Refill Price</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mach Razor Pack Demand</td>
<td>-0.173</td>
<td>-0.061</td>
<td>0.108</td>
<td>-0.089</td>
</tr>
<tr>
<td>Mach Blade Refill Demand</td>
<td>-4.136</td>
<td>2.082</td>
<td>2.942</td>
<td>-0.283</td>
</tr>
<tr>
<td><strong>Club Stores: Mach Blade Refill Price</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mach Razor Pack Demand</td>
<td>-0.017</td>
<td>-0.002</td>
<td>0.071</td>
<td>-0.0002</td>
</tr>
<tr>
<td>Mach Blade Refill Demand</td>
<td>0.818</td>
<td>0.837</td>
<td>-6.800</td>
<td>-0.465</td>
</tr>
</tbody>
</table>

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