MARKETING SCIENCE
Vol. 25, No. 1, January–February 2006, pp. 25–50
ISSN 0732-2399 | EISSN 1526-548X | 06 | 2501 | 0025

DOI 10.1287/mksc.1050.0140

An Empirical Model of Optimal Dynamic Product Launch and Exit Under Demand Uncertainty

Günter J. Hitsch
Graduate School of Business, University of Chicago, 5807 South Woodlawn Avenue,
Chicago, Illinois 60637, guenter.hitsch@chicagogsb.edu

This paper considers the decision problem of a firm that is uncertain about the demand, and hence profitability, of a new product. We develop a model of a decision maker who sequentially learns about the true product profitability from observed product sales. Based on the current information, the decision maker decides whether to scrap the product. Central to this decision problem are sequential information gathering, and the option value of scrapping the product at any point in time. The model predicts the optimal demand for information (e.g., in the form of test marketing), and it predicts how the launch or exit policy depends on the firm's demand uncertainty. Furthermore, it predicts what fraction of newly developed products should be launched on average, and what fraction of these products will "fail," i.e., exit. The model is solved using numerical dynamic programming techniques. We present an application of the model to the case of the U.S. ready-to-eat breakfast cereal industry. Simulations show that the value of reducing uncertainty can be large, and that under higher uncertainty firms should strongly increase the fraction of all new product opportunities launched, even if their point estimate of profits is negative. Alternative, simpler decision rules are shown to lead to large profit losses compared to our method. Finally, we find that the high observed exit rate in the U.S. ready-to-eat cereal industry is optimal and to be expected based on our model.

Key words: new product strategy; product launch; product exit; managerial decision making under uncertainty; Bayesian learning; numerical dynamic programming; dynamic structural models
History: This paper was received July 21, 2003, and was with the author 9 months for 3 revisions; processed by Tülin Erdem.

1. Introduction

New product introductions are common in most markets. Many of these new products "fail," i.e., exit from the market soon after product launch. Urban and Hauser (1993), for example, report an average failure rate of 35% across different industries. In many markets, the failure rate is even higher; in the U.S. ready-to-eat (RTE) breakfast cereal industry, for example, close to 70% of all new products are scrapped within two years after their introduction.[1] These product failures are costly, as the development costs and marketing costs during the launch period can no longer be recovered. A possible explanation for these stylized facts is that firms are often uncertain about the demand and profitability of their new products, and therefore they launch products that are later scrapped. This raises the question of how firms should optimally learn about the true profitability of their products, and how the decision to launch or scrap a product is affected by demand uncertainty. What would be the value of reducing the demand uncertainty, for example, through test marketing? Furthermore, do high exit rates indicate that many firms use suboptimal product launch strategies, or maybe launch products for reasons that are not related to demand uncertainty?

This paper aims to answer these and several related questions. To this end, we develop a model of a decision maker (a firm) who decides whether to launch a new product, and, after product launch, whether the product should be scrapped or stay in the market. The firm is uncertain about the demand, and hence the profitability, of the new product.[2] The firm's uncertainty is formalized through a prior distribution on a parameter that indexes the demand function. Realized sales provide new information about product demand, and allow the firm to update its prior using Bayes' rule. Formally, this is a sequential experimentation problem (Wald 1945), where the decision maker decides period after period whether to

[1] This number is based on the newly launched breakfast cereals during 1988–1992.

[2] Product launch under perfect information is just a special case of the model.

acquire more information in order to test the hypothesis $H_0 =$ "the product is profitable."[3] The decision to keep or scrap the product depends on the costs and benefits of being in the market, i.e., the expected flow of profits and the marketing costs. Therefore, the firm's payoff function has to be based on a realistic demand system, and some important marketing decisions (pricing and advertising) need to be modeled along with the stay-or-scrap decision. Also, we take into account that, based on the state of knowledge about the newly launched product, the equilibrium price of each product in the market may change. The model can be calibrated from actual data, and then used to predict optimal launch behavior and the associated stream of profits.

The model allows us to calculate the expected profit from a product under some level of demand uncertainty. In the spirit of Bayesian decision theory, we can therefore calculate the value from reduced uncertainty in the form of increased expected profits. Together with a cost function of reducing uncertainty, we can then predict the optimal demand for information—for example, in the form of test marketing. The model also shows how, given a point estimate of profits, the decision to launch or scrap a product depends on the degree of uncertainty. For example, should fewer or more products be launched under increased uncertainty? On average, given a level of uncertainty, what is the fraction of products that "fail?" This fraction is the optimal product exit rate, which has to be known before some observed exit rate (such as the 70% exit rate in the RTE cereal industry) can be judged as "too high" or "too low." Our framework is intended to be applied normatively, as a tool for managers, in a wide range of industries. In this paper, we present an application to the specific case of the U.S. RTE breakfast cereal industry. This industry is an interesting example to demonstrate the managerial and economic implications of our model, due to the historically high incidence of product entry and exit.

The questions posed above are all concerned with the relationship between the initial level of demand uncertainty, and the subsequent launch-or-scrap decisions and profit outcomes. Our model assumes that firms behave optimally, conditional on the initial prior. We do not assume, however, that the initial prior itself is chosen optimally. Therefore, we do not contradict the optimality assumption in our model if we ask how subsequent behavior and outcomes change if the initial uncertainty is reduced or increased.

[3] Wald considers a decision maker who does not discount the future, while the decision maker in our model is impatient. See Moscarini and Smith (2001) for a theoretical treatment of optimal experimentation by an impatient decision maker.

The sequential decision problem studied has to be solved using dynamic programming techniques. While generally difficult to solve, simpler decision rules would not correctly address the issues raised above—for example, how the degree of demand uncertainty is related to the launch-or-scrap decision. Using a straightforward but incorrect decision rule, a decision maker would calculate the expected present discounted value of profits, given current information about the product's profitability. The product would then be scrapped if this value was negative or less than a fixed cost of keeping it in the market. This approach, however, neglects that the firm can exercise the option of scrapping the product not only now, but also at any point in future. This option is valuable, because it allows the firm to gather more information about the true product profitability. Exercising this option entails an opportunity cost; therefore, even if the expected value of profits, given current information, is negative, the firm might optimally delay product exit. The option value of a project under uncertainty and irreversibility of actions has been examined in the real-options literature (Dixit 1989, Dixit and Pindyck 1994).[4] Our paper highlights the real-options aspect of a new product launch. In our empirical application, we demonstrate the economic significance of the option, and in particular the resulting value of additional information about the product's profitability.

The use of Bayesian decision theory in marketing research is not new. Churchill (1995) and Lehmann et al. (1998), for example, discuss how to assign an ex ante dollar value to the information to be gained in a test market. Their treatment is mainly focussed on the methodological aspects of Bayesian decision theory, and not on a concrete implementation of these methods. Urban and Hauser (1993) provide an informal discussion of market experimentation during the product launch phase, although only with a focus on the marketing mix, and in particular learning about the effectiveness of advertising.[5] Most closely related to this study is Urban and Katz (1983). They use a Bayesian decision tree to determine whether a product should enter a test market after pretest marketing, and whether it should be launched nationally after a test market. These two decisions are made by comparing the realized market share at each stage to a predetermined cutoff market share. This mechanism

[4] Typically, the real-options literature considers uncertainty about market conditions, while in our application the market conditions are constant, but the decision maker is uncertain about the true state of the world. This problem is closely related to the sequential R&D process examined by Roberts and Weitzman (1981).

[5] Little (1966) and Pekelman and Tse (1980) use adaptive control techniques to show how advertising experiments should be conducted.

maximizes expected profits ex ante, i.e., before the pretest market results are known. However, it is not dynamically optimal, because the decision to launch the product or not is not dependent on the learning in the pretest market. Our model, in contrast, considers sequential learning, and the corresponding value of new information, as a central element of the product launch and testing process.[6] An empirical model of sequential learning has previously been developed and estimated by Erdem and Keane (1996), who consider the case of consumer choice under attribute uncertainty. Methodologically, their paper is closely related to our work.

Our framework builds on the existing insights from Bayesian decision theory and extends the literature on product testing and launch planning in several ways. First, we solve the sequential experimentation problem using numerical dynamic programming techniques. This approach allows us to calculate the optimal path of product exit and stay decisions with respect to the generic objective function of the firm, i.e., the expected PDV of profits. This objective function needs to be derived from a realistic demand system, such as the random coefficients logit model that we employ, which makes an analytic solution to the decision problem infeasible. Second, the solution of a dynamic program ensures that the decisions by the firm are not only optimal ex ante, as in Urban and Katz (1983), but also in any future time period, conditional on what the firm has learned in the market. Third, the numerical solution approach allows us to incorporate intertemporal demand linkages into the model, and then solve for the dynamically optimal marketing mix. Our paper focuses on dynamic advertising, although one could extend the model to allow for dynamic pricing.[7] A managerial application to an actual product launch is feasible using a modern personal computer.[8]

In order to simulate the model and discuss its implications using an actual, empirical application, we first need to estimate the model parameters. We develop a two-step estimation approach. In the first step, most demand parameters are estimated using a generalized method of moments (GMM) approach, due to Berry et al. (1995). This step reduces the computational burden for the second step, where we employ a maximum likelihood (ML) estimator. The ML estimator achieves two goals. First, we can estimate the initial prior variance, i.e., the firms' uncertainty, by matching the observed durations of the products in the market to the durations predicted by the model. Second, we can consistently estimate the intertemporal effect of advertising, controlling for a potential econometric endogeneity problem that arises if firms advertise based on shocks to the advertising effectiveness that are observed by the firms, but not to the researcher. Both tasks are achieved by exploiting the structure of the model; the implementation of the ML estimator therefore requires us to solve the dynamic program (DP) at each evaluation of the likelihood function. Thus, our estimator is a variation of Rust's (1987) nested fixed-point estimator. The estimation problem is much easier for a firm that uses the model to guide its decisions, because it knows its prior and has more knowledge about the effectiveness of its past advertising than a researcher does. Hence, the firm only needs to estimate product demand using the first estimation step in order to examine the normative predictions of the model.

Finally, beyond its main, normative use, the specific application of our model provides a new possible explanation for the observed high product entry and exit rates in the U.S. RTE breakfast cereal industry. Our estimation approach allows us to estimate the initial uncertainty that rationalizes the observed entry and exit decisions. Simulations of the model, given the estimated degree of uncertainty, then reveal the expected, typical outcome in the market. We find that the high observed entry and exit rates can not only be rationalized, but are actually the typical, expected outcome predicted by the model. Previously, some antitrust policy makers and economists have taken the large number of existing and entering products as evidence of product proliferation, i.e., the practice of launching new products to preempt entry by a new competitor (Schmalensee 1978).[9] Our model provides an alternative explanation that should also be considered when judging the performance of the cereal industry. We would like to stress, however, that we are unable to test between the competing hypotheses, and therefore we cannot conclusively rule out that a strategic motive plays a role in the product launch process.

The introduction concludes with a review of the related literature. In §2, we discuss the stylized facts of typical product launches in the U.S. RTE cereal industry. The model is presented in §3. Section 4 provides an

[6] Rao and Winter (1981) consider the related problem of selecting the number of units—e.g., cities—that constitute a whole test market. They also use a Bayesian decision theory framework. Their approach, however, does not allow for sequential decision making.

[7] The focus on advertising is motivated by the particular application to RTE breakfast cereals; an examination of the data reveals that advertising, but not pricing, is set dynamically.

[8] The decision maker needs to estimate or calibrate a demand system, and specify his or her prior about the unknown demand parameter. Using an optimized computer algorithm, the solution to the dynamic decision problem can then be found within a few minutes.

[9] The brand proliferation hypothesis has later been criticized on theoretical grounds (Judd 1985).

overview of our estimation approach and presents the estimation results. In §5, we examine the predictions of the estimated model. Section 6 concludes.

Relationship to the Literature. Numerical dynamic programming has been used in some recent marketing studies. Among the topics examined are consumer choice dynamics in the presence of coupons (Gönül and Srinivasan 1996), consumer choice dynamics in response to price promotions and stockpiling (Erdem et al. 2003, Hendel and Nevo 2005), optimal catalog-mailing policies (Gönül and Shi 1998), and durable product adoption decisions in the presence of quality and price expectations (Melnikov 2001, Song and Chintagunta 2003). Learning models with normal priors have a long history in marketing; Roberts and Urban (1988), for example, examine the adoption of durable goods in the presence of attribute uncertainty.[10] Closest to the formal framework in this paper is the work by Erdem and Keane (1996), who also consider attribute uncertainty. In contrast to Roberts and Urban, they consider consumers as forward-looking decision makers, which gives rise to a dynamic programming problem.[11]

Marketing-mix choices under demand uncertainty have been investigated in the literature on new product development and testing, which we referenced before. Montgomery and Bradlow (1999) examine pricing under uncertainty about the functional form of demand. They, however, do not account for learning. The foundations of the model developed in this paper are related to the literature on monopoly learning. Aghion et al. (1991) provide a general theoretic treatment. McLennan (1984), Trefler (1993), and Raman and Chatterjee (1995) study the pricing decision of a monopolist who does not know its demand curve, Harrington (1995) extends this work to the case of duopoly. Two-sided learning, where both the buyers and the seller(s) of an experience good are uncertain about the product's quality, has been studied in a monopoly setting by Judd and Riordan (1994), and in a duopoly setting by Bergemann and Välimäki (1997). Including consumer learning in the demand side would be an interesting extension to this paper.

In economics, a small but very important literature investigates the competitive effect of entry; see Bresnahan and Reiss (1991) and Berry (1992), for example. Related to this literature, Kadiyali (1996) studies entry deterrence. This literature is only indirectly related to our paper. However, the uncertainty and learning aspect considered here may become important for the economic entry literature once dynamic models are considered; conversely, it would be interesting to extend the work in this paper to allow for strategic interaction, for example, in product introductions.

[10] See Akçura et al. (2004) for an application to consumer learning about drug efficacies.

[11] See also Ackerberg (2003) for a related approach.

2. Product Launches and Product Exit in the U.S. Ready-to-Eat Breakfast Cereal Industry

This section provides a brief overview of the U.S. RTE breakfast cereal industry. We motivate the problem addressed in this paper by some stylized facts and "reduced form" analysis, which suggests that demand uncertainty plays a role during a typical product launch. Also, the stylized facts indicate that advertising may have long-run effects, which motivates a corresponding modeling choice in the next section.[12]

The U.S. RTE breakfast cereal industry provides an ideal social laboratory to study how products are launched and scrapped. First, in this industry product introductions have always been common. Between 1985 and 1992, for example, the major national-brand cereal manufacturers introduced 78 new products. That is, on average, 10 new brands were rolled out every year. Second, the RTE breakfast cereal industry is, relatively speaking, a "simple" industry. That is, complicating dynamic factors such as technological progress, or dynamic price discrimination, as commonly employed over the life cycle of a high-tech durable product, are absent.

Kellogg and General Mills were the market leaders in 1990, followed by Post (a division of Kraft), Quaker Oats, Nabisco, and Ralston Purina.[13] The industry was populated by about 130 national brands. Each product had only a small market share, with the exception of a few big brands such as Kellogg's Corn Flakes (5.2% volume market share), and General Mill's Cheerios (4.6% volume market share).

Mature, established breakfast cereals do not exhibit much time-series variation in their sales, price, and advertising data. Most notably, advertising is considerably more volatile than sales, where the variation in advertising (and sales) occurs around a usually stable brand-specific mean. Newly launched breakfast cereals, on the other hand, exhibit distinct dynamic patterns in their sales and advertising series. Figures 1 and 2 show the national sales, price, and advertising data for two new cereals, Kellogg's Kenmei Rice

[12] A detailed description of the data is provided in §4 and Appendix C is available at http://mktsci.pubs.informs.org.

[13] Two of these manufacturers left the market during the 1990s; Nabisco sold its cereal line to Post in 1993, and Ralston sold its national-brand cereals to General Mills in 1996.

Figure 1 **Kellogg's Kenmei Rice Bran**

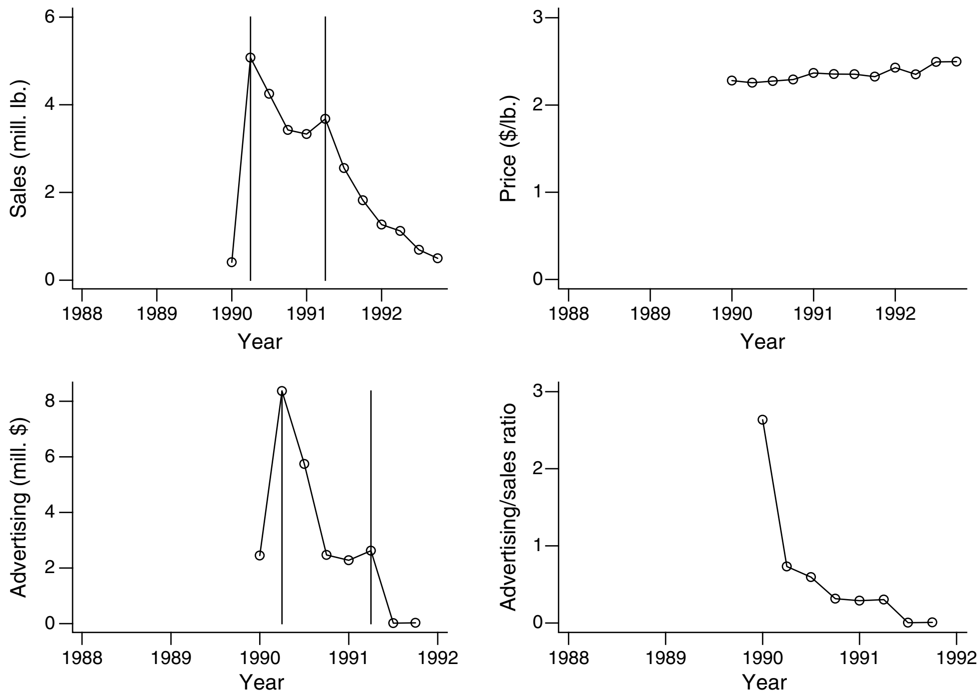
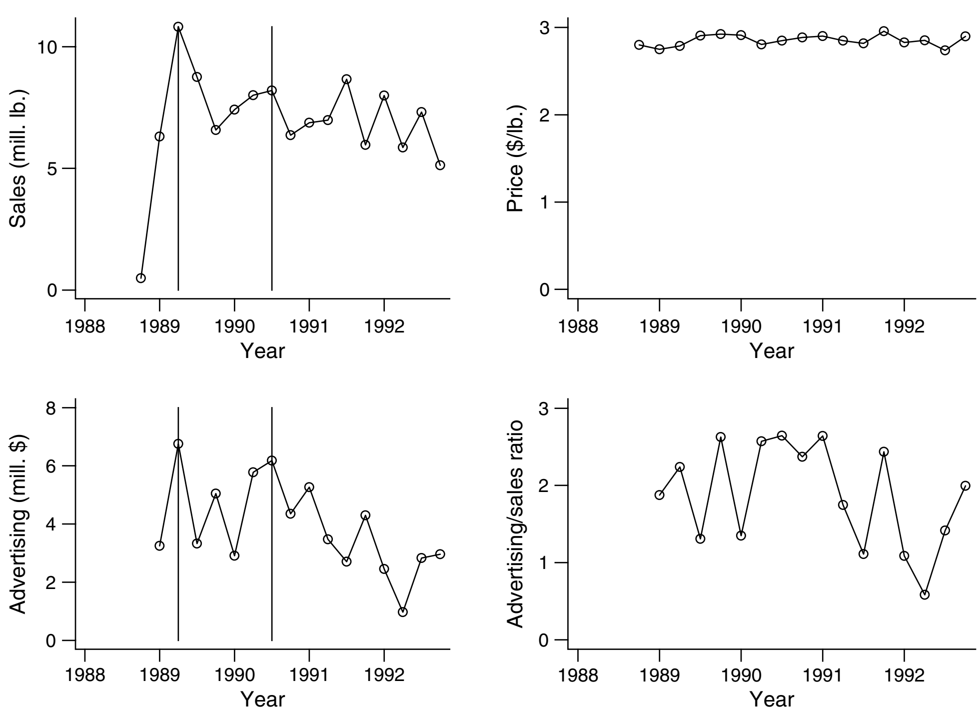


Figure 2 **General Mill's Apple Cinnamon Cheerios**



Bran, a bran-based health cereal, and General Mill's Apple Cinnamon Cheerios. The sales of both cereals peak shortly after product launch.[14] The initial peak in sales is accompanied by an initial peak in advertising. In particular, in both cases the advertising/sales ratio peaks initially and then drops. Pricing, on the other hand, has no apparent dynamic component. These patterns are typical of all new cereal products in our sample.

Only one of the brands depicted, General Mill's Apple Cinnamon Cheerios, was a "successful" product in the sense that it became an established brand. Kenmei Rice Bran, on the other hand, was eventually scrapped. How is this outcome reflected in the data? Most importantly, there is a large difference between the initial advertising/sales ratios, which was about 0.2 for Apple Cinnamon Cheerios and about 2.5 in the case of Kenmei. Both products were supported with similar average advertising spending during the first two quarters after their launch (Kenmei's advertising spending actually peaked at a higher level than Apple Cinnamon Cheerio's spending). However, the realized product sales for Apple Cinnamon Cheerios were much higher than for Kenmei. After the initial peak, Kenmei's advertising spending declined strongly, whereas the decline in product sales was more gradual. The decline in Apple Cinnamon Cheerio's advertising, in contrast, was not pronounced. Table 1 shows that these data patterns are common across all new products used in our empirical analysis.[15] Table 2 presents estimates of a probit model and Cox proportional hazard estimates relating the exit decision to the initial advertising/sales ratio, the total revenue from the product (as a proxy for profitability), and the time in the market. The estimates are based on all new products that were eventually scrapped. The results indicate that the exit probability (hazard) is positively related to the initial advertising/sales ratio, although the effect is not measured precisely.

Note that firms can only set the *ex ante*, not the actually realized advertising/sales ratio. Therefore, a possible interpretation of our findings is that firms are initially uncertain about the true profitability of a product. Initial advertising spending is set to support a potentially profitable product, while sales are a signal of the true profitability. As the firm learns that sales are in fact low, the advertising support is decreased, and eventually the product is scrapped. On the other hand, if realized sales justify the initial ad spending, the firm learns that it has launched a profitable product that stays in the market. Of course, this argument is more suggestive than conclusive. Our approach here is to relate the unobserved uncertainty to a variable that might proxy for it, and then relate this variable to the product exit decision. Below we will use a structural approach to estimate the uncertainty directly from the observed durations until exit.

We observed in the case of Kellogg's Kenmei that sales decline more slowly than advertising. This finding is consistent with intertemporal demand dependencies. If the decay in the impact of advertising on demand is gradual, sales will be higher through past advertising, even if current advertising has been discontinued. As the long-run effect of advertising seems to be an important aspect of the industry studied, we will incorporate it in the model developed in the next section.

Table 1 Summary Statistics

	Scrapped products		Surviving products	
	Year 1	Year 2	Year 1	Year 2
Sales[a]	9.13	4.81	17.86	18.10
Advertising[a]	3.19	0.44	4.16	4.27
Adv/sales ratio	0.37	0.06	0.25	0.24
Peak adv/sales ratio	0.90	0.14	0.39	0.29
Price[b]	2.76	2.75	2.70	2.73

Note. The sample statistics are averages in Years 1 and 2 after product launch.
[a]$ millions. [b]$/lb.

Table 2 Exit Regressions

	(i)[a]	(ii)[a]	(iii)[b]	(iv)[b]
Initial adv/sales ratio	0.769	0.645	0.701	1.202
	(0.516)	(0.806)	(0.693)	(0.883)
Time since launch	0.534	1.492		
	(0.186)	(0.708)		
Revenue		−1.275		
		(0.656)		
Average revenue[c]				−0.235
				(0.174)
Constant	−5.205	−8.548		
	(1.640)	(4.349)		
Observations	57	57	57	57
Log-likelihood	−12.461	−5.486	−8.625	−6.882

Note. The estimates are based on data from all eight products that were eventually scrapped.
[a]Probit model of exit.
[b]Cox proportional hazard model; duration until exit.
[c]Average during first two quarters after launch.

[14] The initial low level of sales is in part due to the fact that the cereals were not launched exactly at the beginning of the calendar quarter. However, all cereals used in our analysis were rolled out simultaneously across all regions.

[15] This sample includes eight products that were eventually scrapped and two cereals that survived. A detailed sample description is given in §4.

3. A Bayesian Learning Model of Dynamically Optimal Product Launch and Product Exit

Overview

We focus on a decision maker (a firm, or a manager) who is uncertain about the demand, and hence the profitability, of a new product. Some characteristics of the product, or, more precisely, the valuations consumers assign to these characteristics, are unknown to the firm. The parameter λ summarizes the effect of these valuations on demand. The firm's uncertainty over λ is expressed in the form of a prior probability distribution $p_t(\lambda)$.[16] The demand uncertainty can be resolved through a product launch, as realized sales provide the decision maker with information about the profitability of the new product. Even after the product launch, there is a potential cost to being in the market, because the product may in fact be unprofitable. The timing of the model is as follows:

1. At the beginning of each period, the firm decides whether to keep or scrap the product.
2. If the firm keeps the product, it sets an advertising level.
3. Prices are determined through a Bertrand pricing game by all firms in the market.
4. Product demand is realized and provides the firm with a noisy signal about the true product quality λ.
5. The firm updates its prior in a Bayesian fashion. The posterior becomes next period's prior, $p_{t+1}(\lambda)$.

Current advertising may impact future demand through a goodwill stock (Arrow and Nerlove 1962), i.e., a specific distributed lag of advertising. This intertemporal demand effect is motivated by the breakfast cereal data used to estimate the model. As discussed in §2, advertising for cereals is more volatile than sales. Hence, *current* advertising can have only a small effect on *current* sales. On the other hand, cereals are advertised heavily. Hence, for such advertising budgets to be optimal, advertising needs to have a long-term impact on sales. Furthermore, we generally observe that sales decline only gradually after advertising for a new, "unsuccessful" cereal is discontinued, which further suggests that current advertising impacts future demand. We allow for this intertemporal demand effect to add realism to the specific application in this paper; it is not, however, a central feature of our model. While we make the advertising decision dynamically optimal, we are unable to allow for strategic interaction in advertising, which poses a severe technical challenge.

A new product launch can be thought of as a statistical test of the hypothesis, $H_0 =$ "the product is profitable." A frequentist approach to this problem would first posit a specific power of the test, and then devise a launch mechanism such that the probability of eventual product failure, i.e., the probability of accepting the null hypothesis while in fact it is false, would be below a certain threshold. In general, this approach will not be optimal with respect to the generic objective function of a firm, the expected present discounted value (PDV) of profits. Our approach, however, is optimal with respect to the PDV of profits by design.

The value of a new product is larger than the simple PDV of profits if there is some demand uncertainty. In a related context, this idea has been explored in the real-options literature (Dixit and Pindyck 1994). The main idea is that the value of the product is derived not only from the PDV of profits, given current information, but also from the option to scrap the product at any point in time. This option allows the firm to delay product exit, an irreversible decision,[17] and therefore to gather more information about the true product profitability. The firm needs to account for this option value when making the decision to stay in the market or scrap the product. This idea can be illustrated with a simple example. Per-period profits from a new product are 1 with probability q, and -1 with probability $1-q$. The simple expected PDV of profits is then given by

$$\frac{1 \cdot q + (-1) \cdot (1-q)}{1-\beta} = \frac{2q-1}{1-\beta},$$

where $0 < \beta < 1$ is the discount factor. According to this objective function, the firm should scrap the new product if q, the probability that profits are positive, is smaller than 0.5. This approach, however, neglects the opportunity cost of exercising the option to scrap the product now. If the firm delays the decision to scrap the product, it gathers more information about the true product profitability. In our simple example, the firm observes profits fully after the first period. Product exit in the second period can then be made contingent on the information acquired in the first period, and the PDV of profits in the first period, assuming dynamically optimal decision making, is given by

$$2q - 1 + \beta \frac{1 \cdot q}{1-\beta}.$$

[16] Firms may reduce some or all of their uncertainty through market research and product testing before the new product launch. This is not inconsistent with our model, because the precision of the prior is a state variable. Hence, a situation where the firm knows the product quality exactly is a special case of the model.

[17] As a more general case, one could allow for partial irreversibility through a sunk exit and start-up cost. This would only be relevant, however, if the firm could obtain additional information about product demand without having to sell the product in the market.

This value is positive if and only if $q > (1-\beta)/(2-\beta)$. For example, if the discount factor is 0.9, which roughly corresponds to an annual interest rate of 10%, the firm should scrap the product only if $q < 0.09$.

The solution of the model considered below involves the same ideas that we just illustrated using a very simple model. The realism needed for an empirical application, however, requires consideration of many details, in particular on the demand function and also on the prior that contains the firm's information. We give a full account of these details below.

Model Details

States and Decisions. The firm's objective is to maximize the expected PDV of the stream of profits generated by a specific product, $\mathbb{E}[\sum_{\tau=t}^{\infty} \beta^{\tau-t}\pi_\tau \mid \mathbf{s}_t]$.[18] The objective function is maximized with respect to a sequence of decision variables $\boldsymbol{\alpha}_t = (E_t, A_t)$, where E_t denotes the exit ($E = 1$) or stay ($E = 0$) decision, and A_t is the current advertising level. These decisions are based on the state vector $\mathbf{s}_t = (h_t, \mu_t, \sigma_t, g_t)$. h_t denotes the age of the product, i.e., the number of time periods passed since the product was launched. μ_t and σ_t index the current prior on the product quality parameter λ. As is explained in more detail below, in our model this prior is a normal distribution, represented by the normal probability density $p_t(\lambda) = p(\lambda; \mu_t, \sigma_t^2)$. g_t is the current "goodwill stock," a measure of the accumulated effect of advertising in demand.

The expected flow of profits in period t is

$$\pi(\mathbf{s}_t, \boldsymbol{\alpha}_t; \lambda) = \int \tilde{\pi}(\mathbf{s}_t, \boldsymbol{\alpha}_t, \boldsymbol{\omega}; \lambda) f_\omega(\boldsymbol{\omega})\, d\boldsymbol{\omega} - A_t - k. \quad (1)$$

$\tilde{\pi}_t$ denotes current profits, gross of advertising, and the per-period cost factor k. k can be thought of as a fixed cost of production, which includes the value of the time managers devote to the specific product, and the opportunity cost of the shelf space in the supermarkets where the product is sold and in the warehouses where it is stored intermittently. The expectation in (1) is taken with respect to the distribution of $\boldsymbol{\omega}$, a vector of demand shocks. The shocks are i.i.d. across time and products, but possibly dependent across geographic markets.

The gross profit function is derived from some underlying demand system. As will be explained in detail below, $\tilde{\pi}_t$ incorporates a vector of equilibrium product prices, which are determined after the current state of the market is realized. In the framework considered here, product prices do not change the transition process of the state vector. Hence, pricing has no dynamic effects on future profit flows, and can therefore be separated from the dynamic decisions, advertising and product exit.

[18] In this formulation, managers are assumed to be risk neutral.

Product Demand. Product demand is modeled as a random coefficients logit demand system (Berry et al. 1995, Nevo 2001). Our approach follows Nevo (2001), who develops and estimates such a demand system for the case of the U.S. RTE breakfast cereal industry. Random coefficients logit models have several desirable properties. In particular, they allow for flexible substitution patterns among differentiated products that can be related to some underlying household heterogeneity structure, yet they can be estimated using market-level data only. Also, the potential econometric endogeneity of prices can be accounted for using instrumental variables. For a detailed discussion, we refer the reader to the aforementioned papers and Nevo's (2000) very accessible "user guide."

More generally, our framework can accommodate any product demand system that has the same underlying latent utility structure as our model. For example, this includes the case of a probit demand system. Even more generally, any demand system with an associated profit function of the form $\tilde{\pi}(\mathbf{s}_t, \boldsymbol{\alpha}_t, \boldsymbol{\omega}_t; \lambda)$ could be employed, although such a formulation may lose the computational tractability of our model.[19]

We consider a market populated by a large number of households who choose among several differentiated products. Household l's indirect utility from product i in time period t is

$$U_{it}^l = \lambda_i - \eta_P^l P_{it} + \mathbf{x}_i' \boldsymbol{\eta}_x^l + \psi(g_{it}, A_{it}) + \tau(h_{it}) + \omega_{it} + \epsilon_{it}^l. \quad (2)$$

λ_i is the quality of product i, i.e., the average valuation consumers assign to all unobserved product attributes. Everything else held constant, the market share of product i increases in λ_i, which is the sense in which the term "quality" should be understood. P_{it} is the product price, and $\mathbf{x}_i$ is a $(K-1)$-vector of observed product attributes that do not vary over time. η_P^l is household l's marginal utility of income, and $\boldsymbol{\eta}_x^l$ expresses household l's preferences for the product attributes in $\mathbf{x}_i$. Both the accumulated effect of advertising in the past and the current flow of advertising have an impact on demand through the function ψ, which is specified in more detail below. $\tau(h_{it})$ is a preference shock that systematically occurs h_{it} periods after product launch. This term is a simple proxy for consumer learning, or variety seeking.

[19] If the firm cannot infer λ from observed sales up to an additive error, the prior and posterior will generally not be conjugate, and hence it will be difficult to represent the prior in a computationally convenient form.

The error term ω_{it} incorporates shocks to the effectiveness of advertising, and taste shocks that are correlated across consumers. Finally, ϵ_{it}^l is an i.i.d. Type I extreme value distributed error term. Households can also choose an outside alternative 0 with (normalized) utility $U_{0t}^l = \epsilon_{0t}^l$.

Following Nevo (2001), we postulate the form

$$\boldsymbol{\eta}^l = \boldsymbol{\eta} + \boldsymbol{\Pi}\mathbf{d}^l + \boldsymbol{\Sigma}\boldsymbol{\nu}^l \tag{3}$$

for the household-specific taste vector $\boldsymbol{\eta}^l \equiv (\eta_P^l, \boldsymbol{\eta}_x^l)$. These preferences are related to $\mathbf{d}^l$, a D-vector describing household l's demographic attributes, and $\boldsymbol{\nu}^l$, a K-vector that captures additional unmeasured sources of taste heterogeneity. $\boldsymbol{\Pi}$ is a $K \times D$ matrix, and $\boldsymbol{\Sigma}$ is a $K \times K$ matrix. $\boldsymbol{\eta}$ is the mean taste vector across all households; hence, $\boldsymbol{\Pi}\mathbf{d}^l$ and $\boldsymbol{\Sigma}\boldsymbol{\nu}^l$ capture the deviation of household l's preferences from this average and relate them to demographics and other unmeasured factors. This random coefficients formulation allows for more realistic substitution patterns than the simple multinomial logit model. In particular, the own price elasticity of a product is related to the price sensitivities of those consumers who have a particular preference for that product, and the cross price elasticity between two products is related to "how close" the products are in attribute space.[20]

We can decompose the indirect utility from each product choice into the *mean utility* component

$$\delta_{it} = \lambda_i - \eta_P P_{it} + \mathbf{x}_i'\boldsymbol{\eta}_x + \psi(g_{it}, A_{it}) + \tau(h_{it}) + \omega_{it}, \tag{4}$$

and the idiosyncractic components

$$v_{it}^l + \epsilon_{it}^l = \tilde{\mathbf{x}}_{it}'(\boldsymbol{\Pi}\mathbf{d}^l + \boldsymbol{\Sigma}\boldsymbol{\nu}^l) + \epsilon_{it}^l. \tag{5}$$

Here, $\tilde{\mathbf{x}}_{it} = [P_{it}, \mathbf{x}_i]$ contains all attributes that have random coefficients. Using these two definitions, the indirect utility from alternative i is $U_{it}^l = \delta_{it} + v_{it}^l + \epsilon_{it}^l$. The population distribution of consumer characteristics in each market is described by the densities f_t^d and f_t^ν. In our empirical application, f_t^d corresponds to the actual distribution of demographics in each market, while $\boldsymbol{\nu}^l$ is a vector of i.i.d. standard normal random variables. $\boldsymbol{\nu}^l$ and $\mathbf{d}^l$ are independent. The market share of product i can be obtained by integrating over household-specific logit choice probabilities:

$$\mathcal{S}_{it} = \frac{1}{\mathcal{M}_t}\int \frac{\exp(\delta_{it} + \tilde{\mathbf{x}}_{it}'(\boldsymbol{\Pi}\mathbf{d} + \boldsymbol{\Sigma}\boldsymbol{\nu}))}{1 + \sum_{j=1}^J \exp(\delta_{jt} + \tilde{\mathbf{x}}_{jt}'(\boldsymbol{\Pi}\mathbf{d} + \boldsymbol{\Sigma}\boldsymbol{\nu}))} \cdot f_t^d(\mathbf{d}) f_t^\nu(\boldsymbol{\nu})\, d(\mathbf{d}, \boldsymbol{\nu}). \tag{6}$$

Here, $\mathcal{M}_t$ denotes the total market size. Equation (6) defines a function that relates the vector of mean utilities to a vector of market shares, $\mathcal{S}_t = (\mathcal{S}_{1t}, \ldots, \mathcal{S}_{Jt}) = T(\boldsymbol{\delta}_t)$.[21] Berry (1994) shows that T can be inverted, such that the vector of mean utilities can be recovered from the observed market shares, $\boldsymbol{\delta}_t = T^{-1}(\mathcal{S}_t)$. This inversion is important both for the estimation of the demand model, and, as we will see further below, for the way the decision maker uses the observed demand to learn about the unknown product quality λ.

The Profit Function. We now derive $\tilde{\pi}_t$, the profit function implied by the demand system. The focus is on a specific product i. F firms compete in the market, and each firm manages one or more products in the set $\mathcal{J}_f$. Let $\xi_{it} \equiv \delta_{it} - \eta_P P_{it}$ be the mean utility net of the price effect (remember that the product price P_{it} is a component of the attribute vector $\mathbf{x}_{it}$). The profit flow of firm f from all the goods in its product line is given by

$$\Pi_f(\boldsymbol{\xi}_t, \mathbf{P}_t) = \mathcal{M}_t \sum_{j \in \mathcal{J}_f} (P_{jt} - c_j)\mathcal{S}_j. \tag{7}$$

Note that all determinants of demand that do not vary period after period are suppressed in the above profit function.

Π_f depends on a vector of mean utilities, $\boldsymbol{\xi}_t$, and thus on the goodwill stocks, advertising levels, and demand shocks of all products in the market. Hence, if firms use all potentially available information about the market, their decisions will be based on a profile of state vectors for each product. Also, they will take into account how these states evolve over time when making their decisions. Finally, they will incorporate their competitors' strategies as a function of the realized states, now and in future, into their own decision making. All these factors together give rise to a dynamic oligopoly model that, given current technology and knowledge, is infeasible to solve with a large number of products. Hence, we make two assumptions to reduce the complexity of the dynamic decision problem. First, we abstract from strategic interaction in advertising. Second, we assume that firms set prices only based on $\bar{\xi}_j = \mathbb{E}\xi_{jt}$, the expectation of ξ_{jt}, for all products j apart from the focal, newly launched product.[22] Otherwise, each ξ_{jt} becomes part of the state vector, and the dynamic programming problem becomes intractable.

[20] A detailed discussion of these issues can be found in the papers cited before.

[21] More precisely, T is also a function of all product attributes $(\tilde{\mathbf{x}}_{1t}, \ldots, \tilde{\mathbf{x}}_{Jt})$, the distribution of household characteristics, and the interactions between household characteristics and product attributes described by $\boldsymbol{\Pi}$ and $\boldsymbol{\Sigma}$.

[22] Once a product has been in the market for some time, the initial conditions, given by the prior distribution p_0 and the initial goodwill stock, no longer impact the distribution of goodwill and advertising. Eventually, all the state and decision variables are distributed according to a stationary, ergodic distribution. In particular, in the long run each ξ_{jt} fluctuates around its mean $\bar{\xi}_j = \mathbb{E}\xi_{jt}$, where the expectation is taken with respect to the ergodic distribution of ξ_{jt}.

The first assumption says that competitors do not change their advertising if a new product enters the market. A direct assessment of how much profits under this assumption differ from true profits is not possible. However, we can provide some indirect evidence by investigating the effect of product entry on equilibrium *prices* and profits. Numerical simulations, based on the parameter estimates to be presented in §4 (Table 5), reveal that a new product launch by some firm f leads to an average price increase of 0.56% among firm f's products, and to an average price decrease of 0.60% among its rivals' products. Furthermore, average per-product profits of firm f, excluding the newly launched product, decrease by 1.78%, while the competitors suffer an average profit loss of 1.90%. If equilibrium prices did not adjust after the new product entry, firm f's average profits would be 0.21% higher, and its rivals' profits would be 0.24% higher compared to profits under the new equilibrium price vector. These findings suggest that profits, calculated without considering a dynamic advertising equilibrium, are only slightly overstated relative to the case where all advertising levels adjust due to strategic behavior. We do not claim that this is a general result; rather, we believe it is specific to a market with a large number of competing products, where a new product entry has only a small impact on the first-order conditions defining an equilibrium price or advertising vector. In other markets, where the equilibrium impact is large, an explicit consideration of dynamic strategic interaction may be warranted.

The second assumption implies that firms do not base their pricing decisions on all current demand shocks. Using numerical simulations, we find that under this assumption median prices are 1.89% higher and median profits are 3.10% higher compared to the case where pricing is based on the realization of all demand shocks.[23]

Under these two assumptions, we can now construct the profit function for each product i. First, substitute all components $j \neq i$ in $\boldsymbol{\xi}_t$ by $\bar{\xi}_j$. Each firm maximizes (7) with respect to $(P_j)_{j \in \mathcal{J}_f}$. The first-order conditions for this problem are

$$\sum_{j \in \mathcal{J}_f} (P_j - c_j) \frac{\partial \mathcal{S}_j}{\partial P_k} + \mathcal{S}_k = 0 \quad \text{for all } k \in \mathcal{J}_f. \tag{8}$$

A Nash equilibrium price vector $\mathbf{P}_t^*$ satisfies the first-order conditions for all firms f simultaneously. Equilibrium profits for firm f are $\Pi_f(\boldsymbol{\xi}_t, \mathbf{P}_t^*)$. Finally, we define the flow of profits from product i as the flow of profits directly attributable to this product, minus the cannibalization cost on the other products in firm f's product line. Eliminate product i from $\mathcal{J}_f$, and recompute the equilibrium profit vector without product i in the market. Denote the resulting equilibrium price vector by $\mathbf{P}_{-i,t}^*$. Product i's profits are then given by

$$\begin{aligned}
&\tilde{\pi}(\mathbf{s}_t, \boldsymbol{\alpha}_t, \boldsymbol{\omega}_t; \lambda) \\
&= \Pi_f(\boldsymbol{\xi}_t, \mathbf{P}_t^*) - \Pi_f(\boldsymbol{\xi}_{-i,t}, \mathbf{P}_{-i,t}^*) \\
&= \mathcal{M}_t \left[(P_{it}^* - c_i)\mathcal{S}_{it} - \sum_{j \in \mathcal{J}_f \setminus \{i\}} [(P_{-i,jt}^* - c_j)\mathcal{S}_{-i,jt} - (P_{jt}^* - c_j)\mathcal{S}_{jt}] \right].
\end{aligned} \tag{9}$$

The term subtracted from product i's direct profits is the *cannibalization effect*.

Advertising. Both past and present advertising have an impact on demand. At the beginning of each period, the firm decides how much to spend on advertising. This flow adds to goodwill g_t, which is a measure of accumulated advertising, and yields a quantity called *added goodwill*, denoted by g_t^a. The added goodwill stock, which is "produced" from goodwill and advertising, impacts mean utility, and hence demand, through the function ψ. Between periods, goodwill depreciates stochastically. Each of these advertising components is implemented through some specific functional form choice. Theory offers little guidance as to these specific choices. Instead, we experimented with different options, and finally chose functional forms that we considered a priori plausible, made the firm's optimization problem well behaved, and yielded a good fit in the estimation procedure.[24]

We first specify ψ, the impact of added goodwill on demand. The multinomial logit demand (as many other discrete-choice models) yields an S-shaped market share function. Therefore, if added goodwill is included linearly in mean utility, the firm will try to capture a large share of the market until diminishing returns to advertising set in.[25] As this outcome is unreasonable, we specify ψ in logarithmic form,

$$\psi(g_t, A_t) = \gamma \log(1 + g_t^a). \tag{10}$$

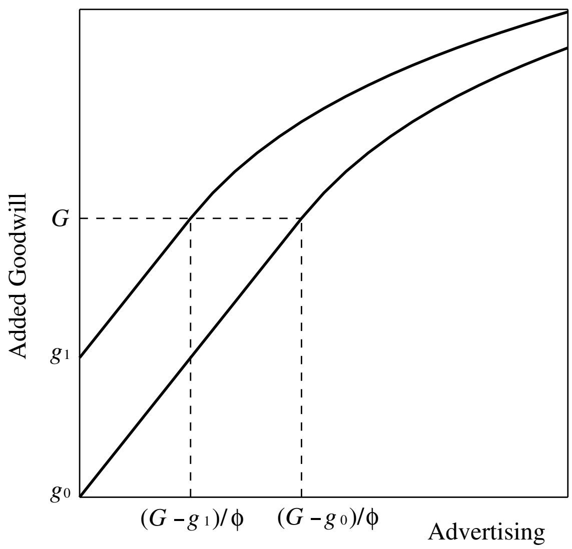
Under this specification the market share is well behaved, i.e., strictly concave in added goodwill, if $\gamma \leq 1$.

[23] Based on means instead of medians, we find a price difference of 3.98% and a profit difference of 4.72%. A visual examination of the distribution of price and profit differences reveals some "outliers," and we thus consider the medians as more representative of the overall price and profit differences.

[24] Hence, we make no claim that our functional form choices are the "correct" ones, or that different functional form choices would not yield a similar solution describing the firm's optimal advertising behavior.

[25] At this point the firm will have captured more than 50% of the market.

Figure 3 **Added Goodwill, "Produced" from *g* and *a***



Added goodwill is created by

$$g_t^a = \begin{cases} g_t + \phi A_t & \text{if } g_t + \phi A_t \le G, \\ G + \log(1 + g_t + \phi A_t - G) & \text{if } g_t + \phi A_t > G. \end{cases} \quad (11)$$

For small goodwill levels ($g_t < G$), the goodwill production function has constant returns, and one dollar of advertising yields ϕ units of goodwill. Beyond the threshold level, G, goodwill production has decreasing returns. This specification includes both the case where goodwill accumulation virtually always has constant returns[26] and the case where diminishing returns already set in for the first dollar of advertising. This specific functional form allows for a strong advertising incentive at small goodwill levels and diminishing returns to advertising at large advertising levels, while remaining simple and dependent on only two parameters. Figure 3 shows the goodwill production function for two different initial goodwill stocks g_0 and g_1.

Added goodwill depreciates between periods, such that the goodwill stock at the beginning of next period is

$$g_{t+1} = \rho_{t+1} \cdot g_t^a. \quad (12)$$

The depreciation rate ρ_{t+1} is stochastic, and assumed to be a lognormally distributed, $\log(\rho_{t+1}) \sim N(\mu_\rho, \sigma_\rho^2)$. Other distributional assumptions could be made; the lognormal distribution, however, is easy to parameterize. We restrict the mean and the variance of ρ such that $\mu_\rho + \sigma_\rho^2/2 < 0$, which implies that $\mathbb{E}(\rho_{t+1}) = \exp(\mu_\rho + \sigma_\rho^2/2) < 1$. Hence, goodwill decays in expectation from one period to the next. Conditional on current goodwill and advertising, goodwill in the next period is also lognormally distributed, $\log(g_{t+1}) \sim N(\log(g_t^a) + \mu_\rho, \sigma_\rho^2)$, and its transition density is denoted by $f(g_{t+1} \mid \mathbf{s}_t, \boldsymbol{\alpha}_t)$.

Note that we can always express the current goodwill stock as a function of an initial goodwill stock, and the sequence of advertising levels and depreciation factors:

$$g_t = \Gamma(g_1, A_1, \ldots, A_{t-1}, \rho_2, \ldots, \rho_t). \quad (13)$$

In this sense, goodwill can be thought of as a nonlinear distributed lag of past advertising.

First-Period Demand Shocks. We include the preference shocks $\tau(h_t)$ in mean utility, which systematically occur during the first T_h periods ($T_h = 2$ in our empirical application) after product launch.[27] These shocks can be due to demand effects such as consumer learning, or variety seeking. The formulation employed allows us to control for these systematic demand shifts after product launch in a reduced-form way.

Beliefs and Learning. The prior distribution $p_t(\lambda)$ describes the firm's belief about the unknown demand parameter λ at the beginning of period t. The observed market shares of all products in a cross section of M markets provide the firm with new information about λ. This information is summarized in a vector $\boldsymbol{\zeta}_t$. Based on $\boldsymbol{\zeta}_t$, the firm updates its prior, $p_{t+1}(\lambda) = p_t(\lambda \mid \boldsymbol{\zeta}_t)$. We assume that at the time of the product launch, the prior is a normal distribution. Under our assumptions on the demand system, the signal $\boldsymbol{\zeta}_t$ has a multivariate normal distribution, and hence the posterior is also normally distributed. Therefore, the priors in each period are described by a normal probability density $p_t(\lambda) = p(\lambda; \mu_t, \sigma_t^2)$, and the firm's uncertainty can be described by the state variables μ_t and σ_t.

Our approach is based on the aforementioned inversion of the product demand system, which allows the firm to infer the mean utility of each product from the observed market shares:

$$T_i(\mathcal{S}_t) = \delta_{it} = \lambda_i - \eta_P P_{it} + \mathbf{x}_i' \boldsymbol{\eta}_x + \psi(g_{it}, A_{it}) + \tau(h_{it}) + \omega_{it}. \quad (14)$$

The firm observes all variables in (14), apart from the product quality λ_i and the demand error ω_{it}; i.e., the firm observes a noisy signal of the product quality, $\zeta_{it} = \lambda_i + \omega_{it}$. The vector $\boldsymbol{\zeta}_t = (\zeta_{it1}, \ldots, \zeta_{itM})$ collects these signals from all M geographic markets. We decompose the error ω_{itm} into a shock that is common across all markets, and a market-specific component,

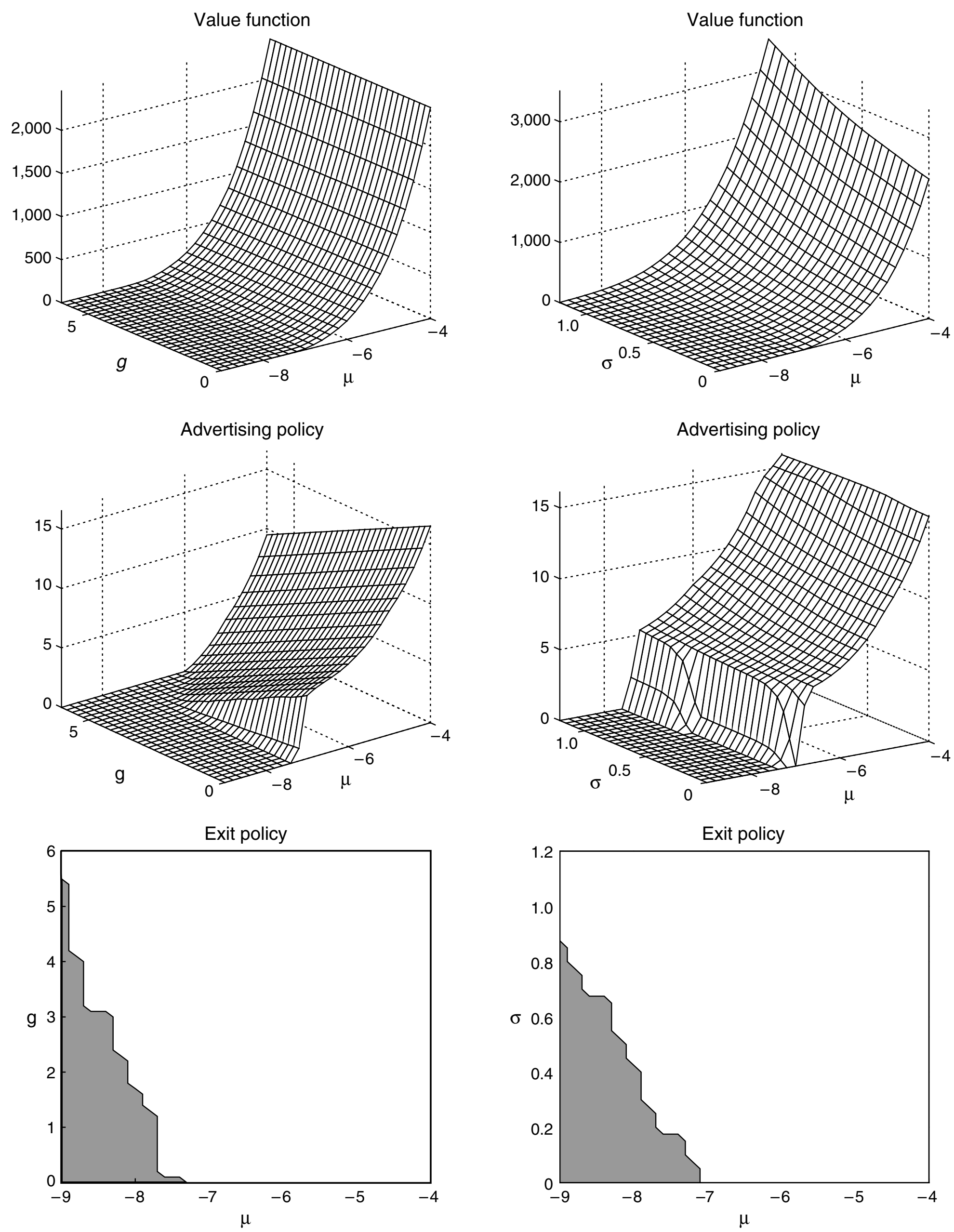
$$\omega_{itm} = \varepsilon_{it} + \nu_{itm}, \quad \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2) \text{ and } \nu_{itm} \sim N(0, \sigma_\nu^2).$$

Shocks that are common across geographic markets can arise from temporal variation of the effectiveness

[26] For a very high goodwill stock the product's market share will be close to 1, and the return to advertising will be smaller than the cost. The firm will therefore not advertise for such goodwill stocks.

[27] That is, $\tau(h) = 0$ for all $h > T_h$.

Figure 4 **Solutions of the DP: The Value Function V, the Optimal Advertising Policy α_A, and the Exit Rule α_E**



Note. The exit region of the state space is shaded gray.

of the advertising copy, for example. Across products and time, the error terms are assumed to be independent. Under these assumptions, the signal $\boldsymbol{\zeta}_t$ has a multivariate normal distribution,

$$\boldsymbol{\zeta}_t \sim N(\mu_t \boldsymbol{\iota}, \boldsymbol{\Omega}_t),$$

$$\boldsymbol{\Omega}_t = \begin{bmatrix} \sigma_t^2+\sigma_\varepsilon^2+\sigma_\nu^2 & \cdots & \sigma_t^2+\sigma_\varepsilon^2 \\ \vdots & \ddots & \vdots \\ \sigma_t^2+\sigma_\varepsilon^2 & \cdots & \sigma_t^2+\sigma_\varepsilon^2+\sigma_\nu^2 \end{bmatrix}.$$

$\boldsymbol{\iota}$ is an M-vector with all components equal to 1. Under Bayesian updating, the posterior $p_t(\lambda \mid \boldsymbol{\zeta}_t)$ has a normal distribution with moments[28]

$$\mu_{t+1} = \mu_t + \sigma_t^2 \boldsymbol{\iota}^T \boldsymbol{\Omega}_t^{-1}(\boldsymbol{\zeta}_t - \mu_t \boldsymbol{\iota}), \quad (15)$$

$$\sigma_{t+1}^2 = \sigma_t^2(1 - \sigma_t^2 \boldsymbol{\iota}^T \boldsymbol{\Omega}_t^{-1} \boldsymbol{\iota}). \quad (16)$$

Equation (15) implies that μ_{t+1}, conditional on the current information summarized in the state and decision vectors, has a normal distribution with mean μ_t and covariance matrix $\mathrm{var}(\mu_{t+1}) = \sigma_t^4 \boldsymbol{\iota}^T \boldsymbol{\Omega}_t^{-1} \boldsymbol{\iota}$. Hence, μ_t has a normal transition density $f(\mu_{t+1} \mid \mathbf{s}_t, \boldsymbol{\alpha}_t) = N(\mu_t, \sigma_t^4 \boldsymbol{\iota}^T \boldsymbol{\Omega}_t^{-1} \boldsymbol{\iota})$.

[28] See the exposition in DeGroot (1969).

Model Solution
The model developed above gives rise to a dynamic programming problem. The solution of this problem is based on the value function, i.e., the expected PDV of profits, which satisfies the Bellman equation

$$V(\mathbf{s}) = \max\Big\{0, \sup_{\boldsymbol{\alpha}_A \geq 0} \mathbb{E}\big[\pi(\mathbf{s}, \boldsymbol{\alpha}; \lambda) + \beta V(\mathbf{s}') \,\big|\, \mathbf{s}, \boldsymbol{\alpha}\big]\Big\}.$$

The numerical details of the model solution are presented in Appendix A.[29]

We now describe the important qualitative aspects of the optimal stay/exit and advertising decisions. Figure 4 shows the value function, advertising policy, and exit rule for a specific parameterized version of the dynamic programming problem. On the left-hand side, the functions are displayed with respect to μ and g, holding σ constant. The right-hand side shows the functions with respect to σ and μ, holding g constant. We are unable to prove that all properties of the depicted solution hold under arbitrary assumptions on the demand system and the other model elements, such as the functional forms by which advertising and goodwill enter demand. However, numerous simulations revealed that for the specific functional forms chosen in this paper, the qualitative aspects of the solution hold under a large variety of chosen parameter vectors.

The Value Function. The value function is increasing in μ, the conditional expectation of the product quality, and goodwill g. This result is obvious. Most importantly, the value of the product is increasing in σ, the uncertainty about the product's profitability. This result reflects the value of the option to scrap the product. This option becomes more valuable when the firm can acquire a large amount of new information about the true product profitability.

Optimal Product Exit. The exit policy, depicted at the bottom of Figure 4, divides the state space into two halves. The exit region of the state space is shaded in gray.

Remark 1. Consider a state vector $\mathbf{s} = (h, \sigma, \mu, g)$ at which the product exits from the market, $\boldsymbol{\alpha}_E(\mathbf{s}) = 1$. Then for any other state vector $\mathbf{s}_0$, such that $h_0 = h$, $\sigma_0 \geq \sigma$, $\mu_0 \leq \mu$, and $g_0 \leq g$, the product will exit, $\boldsymbol{\alpha}_E(\mathbf{s}_0) = 1$.

Under certainty ($\sigma = 0$), $\mu_0 = \lambda$, and all products in the exit region are truly unprofitable products. Holding μ constant, the firm is less likely to scrap a product under increasing uncertainty. This central result is a consequence of the product's option value.

Also, even under certainty, a product is less likely to exit for large values of goodwill g. This is obvious, as goodwill represents an additional profit component.

[29] The appendix to this paper is available at http://mktsci.pubs.informs.org.

Optimal Advertising. Advertising is increasing in μ. This property of the advertising policy implies that product sales and advertising are positively correlated, which is also a stylized fact of the cereal data used in this paper. The positive relationship between advertising and expected sales is closely related to the Dorfman and Steiner (1954) condition, and its dynamic version derived in Arrow and Nerlove (1962). The Dorfman-Steiner condition says that the optimal advertising/sales ratio for a monopolist's product is proportional to the ratio of the advertising elasticity of demand and the price elasticity of demand. The goodwill formulation in this model implies higher marginal revenues from advertising for products with higher product qualities, which implies that advertising is increasing in μ. The positive relationship between advertising and expected sales is not a general property of all advertising models; one can construct examples where it does not hold. However, such an alternative formulation of the advertising technology would not be able to fit the data used in this paper.

The following property of optimal advertising is relevant for the construction of the likelihood function:

Remark 2. Consider the optimal advertising policy at some state vector $\mathbf{s}$ as a function of goodwill g. Then either (i) $\boldsymbol{\alpha}_A = 0$ for all $g \geq 0$, or (ii) advertising is positive and strictly decreasing over the interval $[0, \bar{g})$, and $\boldsymbol{\alpha}_A = 0$ for all $g > \bar{g}$.

Advertising is decreasing in goodwill because of diminishing marginal revenues from goodwill and diminishing returns in added goodwill "production." For sufficiently high levels of goodwill, the marginal return is smaller than the constant marginal cost; hence, no more advertising occurs.[30]

4. Estimation

In this section we present a two-step approach to estimating the parameters of the product launch model. The economic implications of the estimates are then discussed in the next section. Before going into the technical details of the estimators, we will first discuss what the estimation approach ought to accomplish, which depends on the intended use of the model.

[30] Advertising is increasing in the standard deviation of the firm's prior, although quantitatively the relationship is only weak. This relationship is mostly due to the shape of a logit-based demand system. The advertising policy is shown over the range of expected product qualities at which the logit market shares are convex in mean utility. Higher uncertainty then translates into a higher expected value from the product; this, in turn, induces higher optimal advertising. Note that the firm in this model is risk neutral; risk aversion, on the other hand, could imply a negative relationship between advertising and σ.

Consider first the situation of a firm that uses the model to calculate the optimal demand for information, or uses the model to make advertising and exit decisions throughout the launch process. For this purpose, the firm needs to know the demand parameters, and it has to assess the fixed cost k, the discount factor β, and its prior distribution (or uncertainty) about the product quality. In the case of the random coefficients logit model, most demand parameters can be estimated using the "BLP" approach, developed by Berry et al. (1995), which is based on a GMM estimator. Also, if the firm observes the goodwill level of its product, it can treat goodwill as data and directly estimate the goodwill production function (3) and the process governing the evolution of goodwill (12). Furthermore, the impact of goodwill on demand can simply be estimated by including goodwill, or rather a function of it, as a covariate in mean utility.

The estimation approach is more complicated for a researcher who tries to evaluate actual firm behavior. We face two problems. First, one of our objectives is to estimate the firms' degree of uncertainty at the time of product launch. This uncertainty, measured by the prior variance, is not observed, and we therefore need to infer it from observed behavior. Second, the estimation of the advertising parameters is more complicated than in the case where goodwill is observed. Although goodwill can be expressed in terms of the sequence of advertising levels,

$$g_t = \Gamma(g_1, A_1, \ldots, A_{t-1}, \rho_2, \ldots, \rho_t), \tag{17}$$

it is not completely observed by the researcher due to the unobserved initial condition and depreciation shocks ρ. Also, goodwill enters mean utility through the nonlinear functional form $\gamma \log(1+g^a)$. Remember that the logarithm ensures that the objective function of the firm, and hence the solution of the advertising problem, is well behaved. Therefore, an estimation approach based on a moment condition such as

$$\mathbb{E}[Z \cdot \log(1+g^a)] \tag{18}$$

is not feasible, because the value of this expectation at the true parameter value is not known.[31] Furthermore, if the firm observes the realization of the goodwill stock and chooses advertising accordingly (as assumed in the model), we are faced with an econometric endogeneity problem, for then goodwill, an unobserved variable (to the researcher) that directly affects latent utility and hence demand, is correlated with advertising.

To solve these problems, we propose a maximum-likelihood (ML) estimator that yields consistent estimates of both the prior variance and the advertising parameters. The ML estimator is based on the joint likelihood of the sequence of exit decisions, advertising levels, and market shares. The likelihood takes the initial prior, and its effect on the subsequent evolution of firm decisions, explicitly into account. This approach accomplishes two tasks. First, the likelihood incorporates the probability distribution of the duration until a new product is scrapped. Hence, loosely speaking, the ML estimator matches the observed duration to the predicted duration of a product until exit. Because the predicted duration depends on the initial uncertainty, we can infer the prior variance from the data. In order to calculate the probability distribution of the duration of the product in the market, we need to know the optimal exit policy, which, in turn, depends on all model parameters. Hence, we have to solve for the optimal exit policy at each evaluation of the likelihood function. Second, as part of the process to estimate the advertising parameters, we need to derive the distribution of goodwill in each period. As noted before, goodwill can be expressed in terms of the sequence of advertising levels, depreciation shocks, and the initial goodwill stock. Due to the endogeneity problem, however, deriving the distribution of g_t is more complicated than simply applying a change of variables formula. This is because advertising is a function of the current realization of goodwill, which in turn depends on the realization of the depreciation factor ρ_t. Hence, the distribution of $(g_1, \rho_2, \ldots, \rho_t)$, conditional on the observed advertising levels, is different from its unconditional distribution. Our solution to this problem is based on an inversion of the optimal advertising policy, which lets us make an inference about the unobserved goodwill level from the observed advertising decision.[32] Thus, we can calculate the conditional distribution of goodwill in each period. Again, we need to know the firm's optimal policy to calculate the likelihood of the data. Hence, we have to solve the dynamic programming problem at each evaluation of the likelihood function. Therefore, our estimator is an example of a nested fixed-point estimator, first introduced by Rust (1987).

All demand parameters, including the parameters related to the advertising technology, are identified from a sample of mature products only. The standard deviation of the initial prior, on the other hand, can only be identified from a sample of new products, which is obvious given the identification argument given above.

[31] Note that this point would still be valid even if goodwill was observed up to an additive error term.

[32] Ching (2002) faces a similar endogeneity problem. His strategy is to approximate the policy function using a flexible functional form. This function—a polynomial in his application—is then parameterized and estimated jointly with the other model parameters.

Below, we discuss the estimation approach in more detail. In the GMM step, we estimate all demand parameters apart from the parameters related to advertising. All remaining parameters are then recovered in the second, computationally intensive, ML step. This two-step approach eases the computational burden of the estimation procedure, and we can control for price endogeneity in Step 1 using an instrumental variables approach.

Estimation Framework

Step 1: *GMM.* In the first step, we estimate the parameters governing the distribution of the preference parameters $\boldsymbol{\eta}$, and the market- and time-specific fixed effects in demand. We use the BLP approach (Berry et al. 1995), which is based on a GMM estimator. At the heart of this method is an inversion of the market share equation (6). As discussed before, conditional on the observed market shares and the parameters in $\boldsymbol{\theta}_{rc} = (\boldsymbol{\Pi}, \boldsymbol{\Sigma})$, we can invert this equation and retrieve the mean utility vector $\boldsymbol{\delta}_{tm}$ at time t in market m. We then construct the following residual:

$$\tilde{\omega}_{itm}(\boldsymbol{\theta}_{GMM}) \equiv \delta_{itm}(\mathcal{S}_{tm}, \boldsymbol{\theta}_{rc}) - (-\eta_P P_{itm} + b_i + b_t + b_m). \quad (19)$$

Here, b_i, b_t, and b_m are product-, time-, and market-specific fixed effects, and $\boldsymbol{\theta}_{GMM}$ contains η_P, $\boldsymbol{\theta}_{rc}$, and all fixed effects. $\mathbf{Z}_{itm}$ is a vector of instruments such that

$$\mathbb{E}[\mathbf{Z}_{itm} \cdot \tilde{\omega}_{itm}(\boldsymbol{\theta}_{GMM})] = \mathbf{0}. \quad (20)$$

All instruments are stacked into a matrix $\mathbf{Z}$, and all residuals stacked into a vector $\tilde{\boldsymbol{\omega}}$. The GMM estimator, based on the moment conditions in (20), is then defined as

$$\boldsymbol{\theta}_{GMM} = \arg\min_{\boldsymbol{\theta}} \tilde{\boldsymbol{\omega}}(\boldsymbol{\theta})' \mathbf{Z} \boldsymbol{\Phi}^{-1} \mathbf{Z}' \tilde{\boldsymbol{\omega}}(\boldsymbol{\theta}). \quad (21)$$

Here, the weighting matrix $\boldsymbol{\Phi}$ is a consistent estimate of $\mathbb{E}[\mathbf{Z}'\tilde{\boldsymbol{\omega}}\tilde{\boldsymbol{\omega}}'\mathbf{Z}]$.

It is important to note the difference between the residual definition in (19) and the structural error term, ω, that enters mean utility (4). In (19), the time-varying effect of added goodwill on demand is subsumed in the product-specific fixed effect. Still, the moment condition (20) will hold with an appropriate choice of instruments. For our application, we use an average of product prices across regional markets as instruments in the estimation. This approach has been used previously by Hausman (1996) and Nevo (2001). Our choice of instruments could violate (20) if the deviations of added goodwill from its mean are correlated with the market-specific product price. However, our data exhibit only slight time-series variation in product prices. Second, our model predicts that firms hold the added goodwill stock at a roughly constant level, and consequently the predicted product prices are roughly constant through time.

The estimation results allow us to construct the first-order conditions describing the equilibrium market prices. Inverting these first-order conditions, we can recover the marginal cost of production and the profit function of each product. Finally, we estimate the price-adjusted mean utilities ξ_{itm} by

$$\hat{\xi}_{itm} = \delta_{itm}(\mathcal{S}_{tm}, \boldsymbol{\theta}_{rc}) + \hat{\eta}_P P_{itm} - \mathbf{x}_i' \hat{\boldsymbol{\eta}}_x - \hat{b}_t - \hat{b}_m. \quad (22)$$

The price-adjusted mean utilities represent the effect of the product quality parameters, advertising, and age effects on market shares, controlling for price:

$$\xi_{itm} = \lambda_i + \psi(g_{it}, A_{it}) + \tau(h_{it}) + \varepsilon_{it} + \nu_{itm}. \quad (23)$$

In the second estimation step we employ the $\hat{\xi}_{itm}$s to estimate the effect of its components on market shares.

Step 2: *ML.* The remaining parameters are estimated by maximizing the likelihood of the sample $\{\mathbf{y}_{it}\}$. Each data point $\mathbf{y}_{it} = (h_{it}, A_{it}, \boldsymbol{\xi}_{it}, E_{i,t+1})$ contains the age of the product, advertising, the price-adjusted mean utility vector $\boldsymbol{\xi}_{it}$, and the exit decision at the beginning of the next period. Because $\boldsymbol{\xi}_{it}$ is not directly observed, we replace it by the estimate $\hat{\boldsymbol{\xi}}_{it}$. Below, we present an outline of the method used to construct the log-likelihood function. The details of this procedure are laid out in Appendix B, available at http://mktsci.pubs. informs.org.

The main difficulty in constructing the likelihood concerns the distribution of advertising and the exit decision. The densities of these variables depend on the current state, i.e., the prior on the product quality and the accumulated goodwill stock. These state variables are unobserved. However, conditional on an assumption on the initial prior, and the past history of the data, the current prior is uniquely determined. This is because to the econometrician, λ_i is a parameter, and sales (i.e., mean utility) are observed; hence, the econometrician can infer the signals used by the firm to update its priors from Equation (4). Now make goodwill part of the data vector, $\mathbf{z}_{it} = (\mathbf{y}_{it}, \mathbf{g}_{it})$. Conditional on the initial prior, p_{i0}, and the past "data" history, $\mathbf{z}_i^{t-1} = (\mathbf{z}_{i0}, \ldots, \mathbf{z}_{i,t-1})$, we know the current prior, and can then calculate the conditional density of $\mathbf{z}_{it}$. Thus, we can calculate the joint density of the observed data and goodwill stocks. Finally, we find the marginal distribution of $\mathbf{y}_i$ by integrating over the sequence of goodwill stocks and with respect to the initial prior.

The initial prior is a normal probability distribution with mean μ_{i0} and standard deviation σ_0. We assume that μ_{i0} is drawn from $N(\lambda_i, \sigma_0)$, i.e., firms receive an unbiased signal of the true product quality before product launch.

Table 3 Sample Statistics

	Mean	Median	Std. dev.	Min	Max	No. obs.
Volume share (within cereal market), %	1.49	1.01	1.31	0.0005	11.38	45,991
Volume share (within potential market), %	0.28	0.19	0.25	0.0001	1.99	45,991
Price, $/lb	2.69	2.72	0.48	1.1400	4.28	45,991
Advertising, $ mill.	3.16	2.15	2.31	0	10.32	590
Advertising/sales ratio	0.13	0.12	0.17	0	2.63	590

Note. The share and price data are from the full sample used in the first estimation step. The advertising data are from the smaller sample including 32 cereals that are used in the second estimation step.

Data

The model parameters are estimated from IRI[33] scanner data, supplemented by Leading National Advertisers (LNA) advertising data, demographic data from the Current Population Survey (CPS), and product attributes. Summary statistics are provided in Table 3. Below, we provide a brief discussion of the data; further details are given in Appendix C.

The IRI data are at the quarterly level from 1988 to 1992.[34] Sales and prices are observed in a cross section of 47 geographic markets. In order to convert sales into market shares, we calculate the size of the potential market by assuming that each person could consume at most one pound of breakfast cereal per week. The LNA advertising data are only available at the national level. Advertising is measured in dollars, and covers spending on 10 different media types. The demographic data used for the random coefficients estimation are from the CPS March Supplement. Product attributes (sugar and fiber content) are taken from the U.S. Department of Agriculture Nutrient Database for Standard Reference.

The new product entries included in the sample were chosen among all products that entered between 1988 and 1990, such that each product is observed over a sufficiently long time period. Within this subset of products, we eliminated kids' cereals and minor line extension. Kids' cereals differ from other product entries in that the manufacturers often expect only a short lifespan. Furthermore, it is doubtful whether minor line extensions or product modifications are truly new products. After eliminating these types of cereals, we retained a sample of 10 new product entries. Each of these launches received extensive coverage in the trade press, and the associated sales, prices, and advertising spending were clearly identifiable from both the IRI and LNA data sources.

[33] Information Resources, Inc. (IRI) is a Chicago-based market research company.

[34] Due to product entry and exit, some of the products are not observed during all 20 quarters.

For the first estimation step, we employ a sample of 56 cereals. On average, these products account for about 81% of total cereal sales. This sample contains all cereals that have an average category share of at least 0.5%, and all selected product entries. Only 2 of the 10 new products survived beyond 1992. These two "successful" entrants, General Mills' Apple Cinnamon Cheerios and Post's Honey Bunches of Oats, are still sold today. In the second estimation step, we include all new products, and 22 out of the 25 largest established products.[35] A smaller sample avoids some of the heavy computational burden of the ML estimation step. Furthermore, we discarded some products because the associated advertising series revealed some unusual periods of zero advertising. If this is due to measurement error, or institutional features such as budget constraints, the estimated persistence of advertising could be biased upwards.

Estimation Results

Step 1: *GMM.* The GMM estimation approach mostly follows Nevo (2001); hence, the presentation here is intentionally brief. Please see Appendix D, available at http://mktsci.pubs. informs.org. for further results, in particular the first-stage regressions of prices on the instruments, and NLLS logit demand estimates.

As already mentioned, we instrument for the price of product i in market m by taking an average of product i's price in all other regional markets. This yields a set of 20 (one for each quarter) instruments. To allow for heterogeneity in consumer preferences, we interact household income and age with the product price and the sugar and fiber content of each cereal. We also allow for a random coefficient on two category dummies, "all-family" and "wholesome." The main purpose of this specific choice of product attributes is to allow for some meaningful differentiation among the cereals. Many of the newly launched products were health cereals, which are differentiated by their fiber and sugar content. Our model specification thus allows for different substitution patterns across the different cereal categories.[36]

The random coefficients logit parameter estimates are shown in Table 4. The model also includes product-, market-, and time-specific fixed effects (these estimates are not displayed).[37] The results are

[35] These 25 largest products coincide with the sample used by Nevo (2001).

[36] We estimated several a priori reasonable models, and chose the final specification based on the GMM objective function and the individual statistical significance of each parameter.

[37] Including advertising in the model leaves the parameter estimates virtually unchanged. This remains true if lagged values of

Table 4 Random Coefficients Logit Estimates[a]

	Means[b]	Standard deviations	Interactions with demographic variables: log(income)	Interactions with demographic variables: $\log(1+\text{age})$
Price	−4.903		2.200	2.594
	(2.579)		(0.668)	(2.968)
Sugar	0.420		−0.624	
	(0.014)		(0.224)	
Fiber	−3.808			8.783
	(0.100)			(5.707)
All-family	17.249	11.779		
	(0.351)	(3.737)		
Wholesome	3.821	4.916		
	(0.180)	(2.494)		
GMM objective	21.78			
No. obs.	45,991			
% Price coefficients > 0	1.21			
Median margin (%)[c]	43.70			

[a]Product-, market-, and time-specific fixed effects were included in the model. Standard errors are in parentheses.
[b]Estimates from a minimum-distance procedure.
[c]The margins were calculated as $(P-c)/c$.

Table 5 ML Parameter Estimates

	(i) Estimate	(i) Std. error	(ii) Estimate	(ii) Std. error
Goodwill accumulation				
ϕ	0.795	0.004	0.620	0.007
G	4.608	0.006	4.334	0.019
Advertising effect γ	2.818	0.003	2.674	0.007
Goodwill evolution				
$\exp(\mu_\rho)$	0.567	0.007	0.613	0.0002
σ_ρ	0.889	0.006	0.585	0.006
Demand shocks				
σ_ε	0.588	0.007	0.547	0.007
σ_ν	0.891	0.000	0.795	0.0002
Initial demand effects				
$\tau(0)$	0.420	0.006	0.416	0.001
$\tau(1)$	0.463	0.007	0.594	0.010
Fixed cost (\$ million) k	1.393	0.385	0.750	(calibrated)
Std. dev. of initial prior σ_0	1.311	0.005	1.041	0.0001
Discount factor β	0.975	(calibrated)	0.975	(calibrated)
Mean PDV, all products	147.14			
Std. of PDV, all products	151.27			
Mean PDV, new products	31.45			
Std. of PDV, new products	98.72			
No. obs. (shares)	25,006			
No. obs. (advertising)	590			

Note. The estimates are based on data from 22 established and 10 new products. In model (ii), the fixed-cost parameter k was calibrated at \$0.75 million. For each model we also estimated the product-specific quality parameters λ_i, which are not displayed.

intuitive: High income and older consumers are less price sensitive than the average household. Also, older consumers have a relatively strong preference for cereals that are high in fiber content, while high-income consumers prefer cereals that are low in sugar. These results are reflected in the estimated substitution patterns. While the median price elasticity across all cereals is −3.30, health cereals, in particular, are characterized by significantly lower elasticities.[38]

The demand estimates are used to recover the marginal production cost values c_i from the first-order conditions (8). The implied median gross margin is 43.7%. This value is similar to the findings from previous research. (Cotterill 1999 reports an average gross margin of 46.0%; Nevo's 2001 main estimate is 42.2%.)

Based on the estimates and the inferred cost values, we can calculate the gross profit functions $\tilde{\pi}$, as given in Equation (9). Note that the profit functions and first-order conditions are expressed in terms of retail prices. However, our focus is on the decisions made by the cereal manufacturers, who have only direct control of the wholesale price.[39] We assume that the average retail markup in the breakfast cereal industry is 18%. This number is consistent with the trade literature and data on retail and wholesales prices from the Dominick's Finer Foods database at the University of Chicago.[40] Under this assumption on the retail markup, there is a fixed relationship between the retail price P and the wholesale price P_w ($P = 1.18 \cdot P_w$), and the first-order conditions (8) can be expressed directly in terms of the wholesale prices.

Step 2: *ML.* The main results from the ML estimation step are displayed in Table 5, column (i). The implications of these estimates will become clear in the next section, where we simulate the calibrated model. Here we only highlight some key findings. We did not attempt to estimate the discount factor β, but instead chose a calibrated value of 0.975, which

advertising are added. We take this as additional evidence that our estimates of the price-adjusted mean utilities ξ_{itm} are consistent (see the discussion in the previous subsection).

[38] For example, Heartwise, a newly launched health cereal, has an estimated price elasticity of only −1.90; Special K's estimated price elasticity is −1.84.

[39] Our data also do not have enough detail to examine whether the incentives of the retailer and manufacturer to scrap or keep a product coincide or not. Conversations with industry observers suggest that manufacturers such as Kellogg and General Mills have much channel power, and hence we do not expect that some retailers can "force" these firms to abandon a new product sooner than privately optimal for them.

[40] See http://gsbwww.uchicago.edu/research/mkt/Databases/DFF/DFF.html.

corresponds to an annual interest rate of 10.7%. All parameters are estimated precisely.[41]

The estimated standard deviation of the initial prior is 1.311 and indicates substantial demand and profit uncertainty at the time of product launch. In fact, at small values of σ_0 the likelihood of the data is 0, because the model cannot explain why some of the unprofitable products were launched. The effect of advertising is highly persistent; the median carry-over of current goodwill into the next period is 56.7%. For a comparison, Appendix D presents logit estimates with current and four lagged values of advertising included. Lagged advertising has a statistically significant effect, but the overall effect size is much smaller than estimated here. This provides support for the structural endogeneity correction provided by the ML estimator. The demand shocks are correlated across geographic markets, as indicated by the standard error of the common variance component, $\hat{\sigma}_\varepsilon = 0.588$. Hence, even though firms receive M signals about the product quality in each period, learning is impeded due to the correlation of signals across markets.

The estimate of the fixed-cost parameter k is \$1.393 million (per quarter). The standard error of k is much higher compared to the standard error of the other parameters, which indicates that k is hard to identify precisely. An identification problem arises, for example, if k is heterogeneous across firms (or products), or even varies across time periods. We therefore checked the sensitivity of the results by re-estimating the model for different calibrated values of k chosen on a grid, $k = 0, 0.25, 0.5, \ldots$ million. As one example, column (ii) in Table 5 shows the parameter estimates for $k = \$0.75$ million, which is the smallest fixed-cost value examined at which the likelihood of the data was positive. Overall, we find that the parameter estimates do not change much, and in particular the main predictions of the model, discussed in §5, remain mostly unchanged (see Appendix E, available at http://mktsci.pubs.informs.org.).

A product-quality parameter, λ_i, is estimated for each product. These estimates are not interesting per se; instead, Table 5 presents the implied distribution of expected profits. Among the new products, the average expected PDV of profits is \$31.45 million and has a standard deviation of \$98.72 million. The large standard deviation relative to the mean reflects that the majority of new products in our sample has a value of zero, while some of them are profitable.

5. Model Predictions

We now return to the questions posed in the introduction. First, we show how our model can be used to calculate the dollar value of improved information about product profitability. This value directly implies the optimal demand for information, in the form of market research. Second, we investigate how the degree of demand uncertainty influences a rational firm's willingness to launch new products and, at the same time, "tolerate" product failures. In particular, should a firm be more, or less, willing to launch a new product, holding the point estimate of expected profits constant, if the firm is more uncertain about the true product profitability? We also illustrate the value of our model relative to other, possibly simpler, frameworks. In the introduction we stressed the importance of the option of scrapping the product now or in the future, which is valuable because it allows the firm to gather more information about product demand. We demonstrate the quantitative significance of the value of additional information by comparing our model to a simpler framework, where the product launch and exit decisions are based purely on an estimate of expected future profits. Second, we consider an alternative decision rule, where the decision maker launches a product only if the "failure" probability is below a certain level. A manager whose career prospects are linked to the number of "successful" launches or the number of product failures might follow such a decision rule.

We explore all these issues using the empirical implementation of our model for the specific case of the U.S. RTE breakfast cereal industry. A limitation of this approach is that the results cannot be simply generalized to any other market. On the other hand, our specific illustration of the more general concepts allows us to discuss the magnitudes and economic importance of these concepts in an industry where product launches are important.

The results presented are obtained from counterfactual simulations, where either the initial uncertainty of the decision maker, or a specific behavioral assumption about the firm's decision rule is varied. This approach is not inconsistent with our model or the estimation strategy, where we assumed that the data were generated from the actions of rational firms. Both the model and the estimation results are conditional on some initial prior uncertainty, and we do not assume that this initial uncertainty itself is optimal. Hence, it is not contradictory to investigate how the outcome of the launch process changes under some different initial degree of uncertainty, or under some alternative, suboptimal decision rule.

Preliminaries. We first clarify some concepts discussed in this section. The expected profit from a product with true quality λ, and a decision maker's prior $p_0(\cdot; \mu, \sigma^2)$, is defined as

$$\Pi(\lambda, \mu, \sigma) = \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \pi(\mathbf{s}_t, \boldsymbol{\alpha}_t; \lambda)\right]. \qquad (24)$$

[41] The first step estimation error does not affect the asymptotic covariance matrix of the step two estimator. See Appendix B for a detailed explanation.

The dependence of expected profits on λ is obvious. Furthermore, the prior information on λ influences profits through the sequence of stay/exit and advertising decisions taken through time. Note the difference between this profit measure and the value function of the firm,

$$V_t = \int \Pi(\lambda, \mu, \sigma) p_t(\lambda; \mu_t, \sigma_t^2)\, d\lambda. \tag{25}$$

The value function measures expected profits with respect to the firm's prior, while Π measures expected profits from the point of view of an observer who knows both the true product quality and the prior of the firm.

A profitable product is such that $\Pi(\lambda, \lambda, 0) > 0$. That is, if the firm is certain about the product quality λ, and makes its decisions accordingly, average realized profits are positive. For an unprofitable product, $\Pi(\lambda, \lambda, 0) < 0$. If there was no uncertainty, the product launch process would be trivial, and only profitable products would be launched. However, under imperfect information, firms receive only a noisy signal about the true product quality. By assumption, $\mu \sim N(\lambda, \sigma^2)$, such that the initial prior is also normal, i.e., $p_0 = N(\mu, \sigma^2)$. Imperfect information leads to "wrong" decisions. For example, a sufficiently negative draw of μ can lead the firm to scrap a profitable product. A sufficiently positive draw can lead the firm to launch an unprofitable product. Finally, advertising will be distorted relative to its level under perfect information.

Suppose a firm develops new products with quality λ, where λ is drawn from a distribution with density f_λ. Under this development process, the average loss from imperfect information can be measured by

$$\begin{aligned} L(\sigma) &= -\mathbb{E}[\Pi(\lambda, \mu, \sigma)] \\ &= -\iint \Pi(\lambda, \mu, \sigma) f_\mu(\mu; \lambda, \sigma^2) f_\lambda(\lambda)\, d\mu\, d\lambda. \end{aligned} \tag{26}$$

The loss is minimized at $\sigma = 0$. The loss function depends on the distribution of λ, i.e., the nature of the firm's product development process, in two ways. First, f_λ determines the scale of profits. Second, given a degree of uncertainty, the loss function will increase in the scope for "mistakes." For example, if the firm is faced with both profitable and unprofitable products with equal probability, it will often launch an unprofitable product or fail to launch a profitable product. On the other hand, if the firm always knows whether the product is profitable or unprofitable, even though the exact level of profits is unknown, there is no scope for mistakes and hence no loss in profits due to uncertainty.

It is important to understand how the loss function differs from the value function of the firm. The value function is also an expectation of future profits; however, the expectation is taken with respect to the subjective information contained in the prior. Hence, the subjective value from the product may increase in the standard deviation σ, because a higher standard deviation may increase the probability that the new product opportunity is profitable. The loss function, however, increases in the standard deviation because the incidence of wrong stay/exit and advertising decisions increases. The distinction between the loss function and the value function is of utmost importance for managerial decision making. A firm should expect that its average profit from a new product opportunity decreases under a higher degree of uncertainty. However, because the value of the product increases in the uncertainty about profits, the firm should be *more* willing to launch a new product.

We calculate the loss function through a simulation process. First, a "product opportunity" is drawn. A product opportunity is determined by the quality parameter λ, the production cost c, and the product attributes that determine the substitution patterns incorporated in the profit function.[42] Consistent with the learning model, μ_0 is drawn from $N(\lambda, \sigma_0^2)$. We then simulate the optimal advertising and exit decisions through time, and record the resulting profits.[43] Finally, we calculate the resulting PDV of profits. The same simulation process also allows us to calculate the percentage of all product opportunities that are launched, and the percentage of all launched products that eventually exits. The latter is called the "optimal expected exit rate."

The Value of Improved Information. We first investigate by how much the expected profit from a product opportunity increases under improved information, i.e., under lower uncertainty about the product quality. From this value we can derive the optimal demand for information, for example, in the form of more or better market research. Formally, if $C(\sigma)$ represents the per-product cost of obtaining a prior with precision $1/\sigma^2$, the firm should reduce its uncertainty to the level

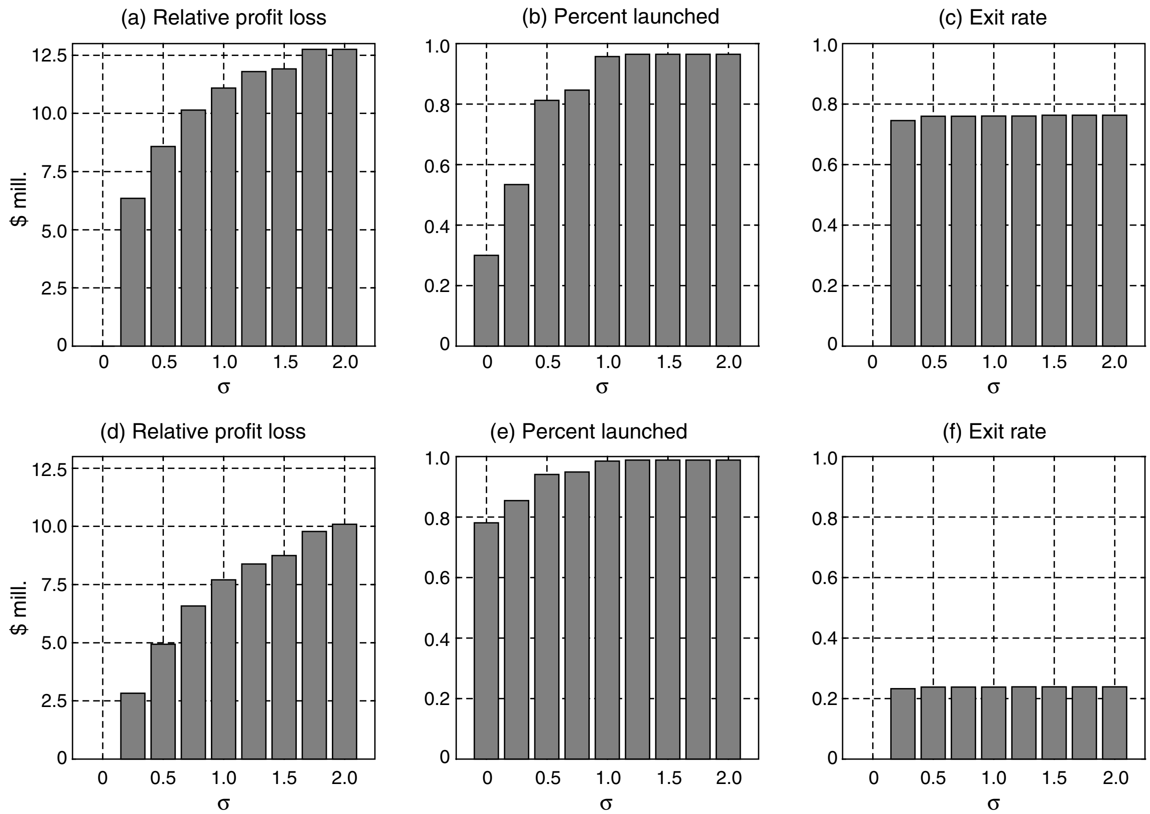
$$\sigma^* = \arg\max_{\sigma \geq 0} \{-L(\sigma) - C(\sigma)\}. \tag{27}$$

We make two different assumptions on the distribution from which a new product opportunity is drawn. First, we assume that the products are drawn from Λ_{new}, the empirical distribution of all newly launched products. Second, we assume that the products are

[42] Note that in the definition of the loss function above, we implicitly assumed that products are only distinguished by their product-quality parameter. The extension to differences in production costs and substitution patterns is only a matter of changing the notation.

[43] For each product draw, we simulate 20 time periods.

Figure 5 Top Row: Products Are Drawn from the Empirical Distribution of All Newly Launched Products. Bottom Row: Products Are Drawn from the Empirical Distribution of All Products in the Sample



drawn from Λ_{all}, the empirical distribution of all products, new and established, in our sample. The first distribution contains a larger fraction of unprofitable products than the second (Table 5).

Figures 5(a) and (d) show $L(\sigma_0) - L(0)$, the profit loss under a specific σ_0 (the standard deviation of the prior at the time of product launch) relative to the baseline profit under perfect information.[44] The profit loss can be large; at the estimated standard deviation $\sigma_0 = 1.31$, it is \$11.8 million under Λ_{new} and \$8.4 million under Λ_{all}. As a percentage of the expected profit under perfect information, the loss is 37.5% under Λ_{new} and 5.7% under Λ_{all}. The reason for this difference is straightforward. Under the empirical distribution of all newly launched products the incidence of unprofitable products is higher than under the empirical distribution of all products in the market. Hence, the average profitability of a new product is lower (\$31.45 million versus \$147.14 million), and the scope for making the wrong entry or exit decision is higher. Hence, relative to average profits under certainty, the value of information is higher under Λ_{new}.

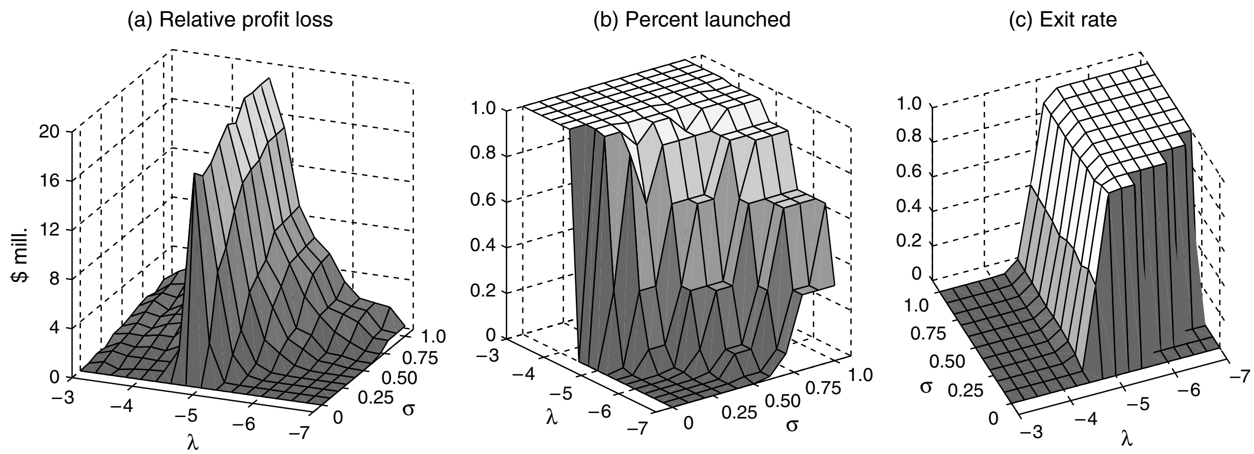
Optimal Entry and Exit Rates. We next turn to the question of how firms should change the fraction of new product opportunities launched under increasing demand uncertainty. Furthermore, we calculate the optimal exit rate. Figure 5 displays the percentage of new product opportunities launched and the product exit rate under Λ_{new} and Λ_{all}.[45] Under full information, 30% of all product opportunities are launched under Λ_{new}, and 78% of product opportunities are launched under Λ_{all}.[46] The difference in the launch rates reflects the underlying occurrence of profitable products under the two distributions. The percentage of products launched increases strongly in the degree of uncertainty. If products are drawn from Λ_{new}, 96.5% of all product opportunities should optimally enter the market at the estimated uncertainty, $\sigma_0 = 1.31$. Even under much smaller levels of uncertainty, more than 80% of all product opportunities should be launched. Consequently, firms incur large product exit rates; for example, under Λ_{new} close to 80% of all newly launched products exit if the initial uncertainty σ_0 is larger than 0.25. Hence, optimal, forward-looking behavior can imply that under even

[44] All profit numbers shown are in 2004 dollars.

[45] The exit rate is defined relative to all products that were actually launched, i.e., not including product opportunities that never entered the market.

[46] In other words, 30% and 78% of all new products drawn from these two distributions are profitable.

Figure 6 Expected Profit Loss, Entry, and Exit Under Different Product Qualities



small degrees of uncertainty, a large fraction of all possible new products should be tested in the market. Furthermore, in our example a high incidence of product exits is fully consistent with optimal decision making.

Profit Loss, Entry, and Exit Given a Product Quality λ**.** Clearly, the expected profit loss, and the product entry and exit rates, are dependent on the distribution of product qualities and hence profits. Figure 6 clarifies these dependencies. The expected loss, entry rate, and exit rate are depicted conditional on the true product quality λ, and a degree of uncertainty σ_0.[47] As λ increases, products become more profitable. In this specific example, there is a profitability threshold $\bar{\lambda} \approx -4.25$ such that products with quality $\lambda > \bar{\lambda}$ are profitable, and products with quality $\lambda \leq \bar{\lambda}$ are unprofitable. Under perfect information, firms select only the profitable products. For example, if $\lambda = -5.5$, the new product opportunity is not launched in the market. Under imperfect information, however, product entry rises strongly in σ_0, holding λ constant. For example, if $\lambda = -5.5$ and $\sigma_0 = 1$, the product entry rate is 90.8%. This is despite the fact that more than 89% of the μ_0 draws are below the entry threshold $\bar{\lambda}$ (remember that $\mu_0 \sim N(\lambda, \sigma_0)$). That is, with probability 0.89 the firm's best guess of the product quality parameter λ is such that the product is unprofitable, but nonetheless the product is launched. In this example, optimal market experimentation leads firms to minimize Type I errors under the null hypothesis $H_0 =$ "the product is profitable." Profitable products are mostly correctly classified, and firms would rather launch an unprofitable product than miss a profitable one. As a consequence of this rational selection process, we observe a large incidence of Type II errors, i.e., product exits.

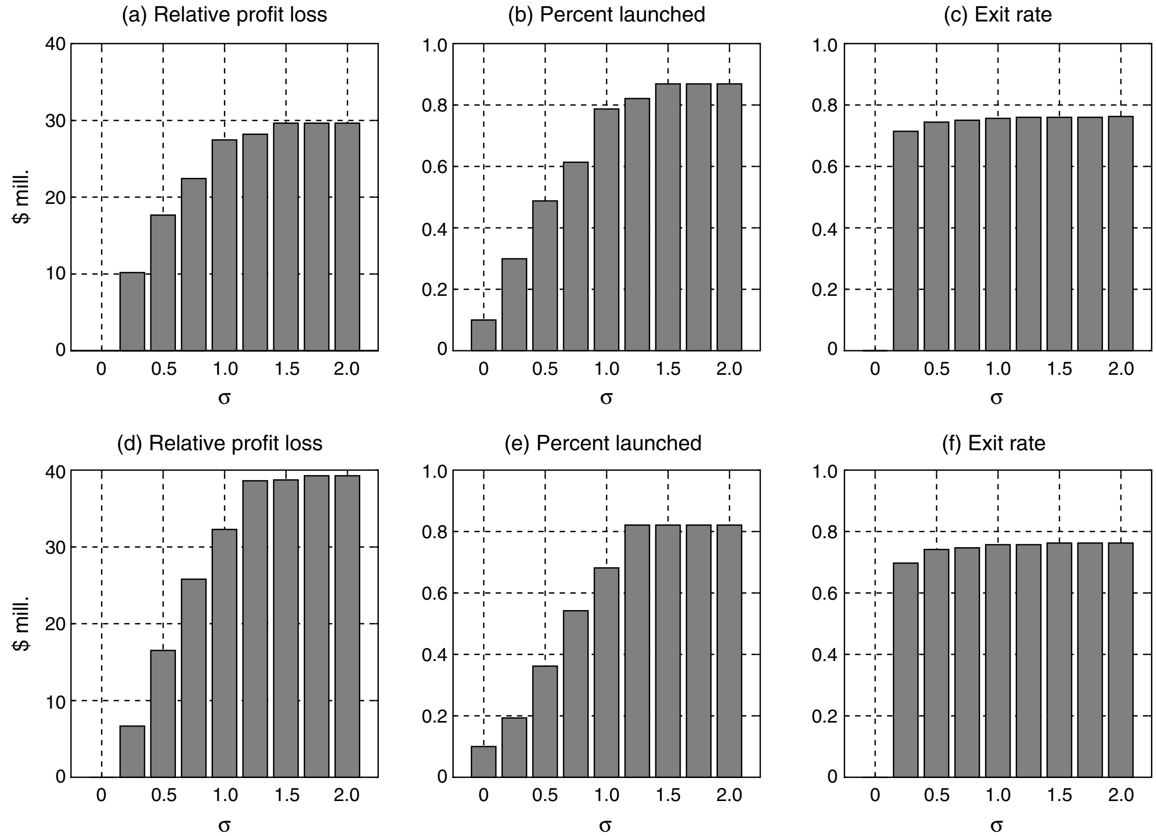
Figure 6(a) shows that the profit-loss peaks at values of λ that are slightly below the profitability threshold. The value of information will therefore be highest if product opportunities are drawn from a distribution of qualities centered around the profitability threshold.

Effect of a Fixed Cost of Entry. The previous experiments were performed under the implicit assumption that the product had already been developed, and, accordingly, that the firm did not have to incur an entry cost. We now investigate how the results change if the cost of entry is positive. This counterfactual experiment corresponds to the situation where the firm is uncertain about the true product profitability, but has to incur some sunk cost before it can start collecting information in the market. These costs can include product development costs, costs to develop a distribution channel, etc. The decision process of the firm has to be modified under this scenario. If the expected value of the new product is larger than the sunk cost of entry, the product enters the market. Otherwise, the fixed cost is not incurred, and the product is never launched.

Figure 7 shows the loss function and the entry and exit rates for a fixed cost $10 million (top row) and $20 million (bottom row). Such development costs are plausible in the case of the cereal industry. Product opportunities are drawn from Λ_{new}. As expected, the value of information is higher for a larger fixed cost of entry, holding the degree of demand uncertainty constant. For example, at $\sigma_0 = 1.31$, the expected profit loss is about $28 million if the entry cost is $10 million, and about $39 million for an entry cost $20 million. The firm's willingness to launch new products is more "conservative" relative to the case of no fixed entry cost, although the change is less dramatic than

[47] The numbers shown are for Kellogg's Heartwise, a new health cereal that was eventually scrapped.

Figure 7 **Product Types Are Drawn from Empirical Distribution of All Newly Launched Products**



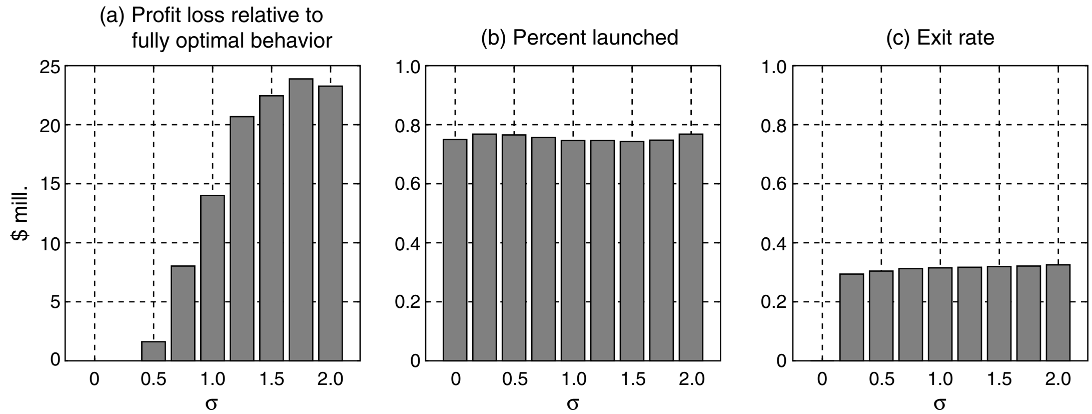
Note. The fixed cost of entry is \$10 mill. (top row) and \$20 mill. (bottom row).

one might expect. Comparing Figures 5 and 7, we see that the percent of products launched is still strongly increasing in the uncertainty σ. At $\sigma_0 = 1.31$, the percentage of launched products is about 82% for both a fixed cost of \$10 million and \$20 million, and the concomitant exit rates are approximately 76%.

Interpreting the Stylized Facts in the U.S. Cereal Industry. The previous findings have strong implications for our interpretation of the observed stylized facts—high entry and exit rates—in the U.S. RTE breakfast cereal industry. Indeed, we find that at the estimated degree of uncertainty, the observed entry and exit rates are as expected from our model of fully optimal decision making. It is important to note that we do more than just rationalize the data: We show that the stylized facts correspond to typical, *expected* behavior. A simple way to rationalize the launch of an unprofitable product is as follows. Assume that the firm's initial prior on λ is such that $\mu_0 > \lambda$ is above the profitability threshold. Then, we would say the product was launched because the firm overestimated the true profitability. We could invoke this argument for all unprofitable products launched. Of course, this would be implausible under rational behavior, although possible in principle (and thus have a positive likelihood). However, the high rate of product launches (Table 5b) implies that even if the firm has a prior such that μ_0 is somewhat below λ, it will still launch the product under the estimated degree of uncertainty. Therefore, the observed entry rates and concomitant product failures are to be expected if the cereal manufacturers are as rational as the firms in our model. Of course, this argument is conditional on the estimated degree of uncertainty; given that we do not observe the cost of reducing uncertainty below this level, we are unable to make any statements about the optimality of the initial uncertainty choice.

Suboptimal Behavior: Disregarding the Value of Additional Information. We previously stressed the importance of the sequential experimentation aspect of the product launch problem. Optimal dynamic decision making is forward looking, and takes into account that more information about the true product quality can be obtained if the product is launched or stays in the market. We contrast the fully optimal behavior to a simpler form of decision making that is focused on a forecast of profits given *current* information only. This alternative decision process does not take into account the option value of scrapping

Figure 8 **Suboptimal Decision Making: The Value of Additonal Information by Delaying Exit Is Not Taken into Account**



Note. Products are drawn from the empirical distribution of all products in the sample.

the product at any point in the future, which allows the firm to gather additional information by delaying exit. In order to demonstrate the economic importance of this additional information, we recompute the decision problem and assume that the decision maker does not anticipate the future change of his prior when the realized sales provide new information about the product's profitability. That is, instead of anticipating that the belief on the product quality tomorrow, μ_{t+1}, can change in response to the information provided by product sales, the decision maker acts as if $\mu_{t+1} = \mu_t$. On the other hand, we retain the assumption that advertising is set in a fully optimal, forward-looking fashion.

Figure 8(a) displays the difference in profits between fully optimal behavior and the suboptimal decision rule described before. The value of additional information and the option value of delay can be large. For example, at $\sigma_0 = 1.31$, their combined value is about \$20.7 million. What is the major source of the profit loss under suboptimal behavior? Comparing Figures 5(b) and 8(b), we see that relative to fully optimal behavior, the percent of product opportunities launched hardly increases in the amount of uncertainty. Hence, many profitable product opportunities are foregone. As the firm does not recognize the full value of the product, it now tolerates a larger fraction of Type I errors, i.e., rejects too many profitable products.

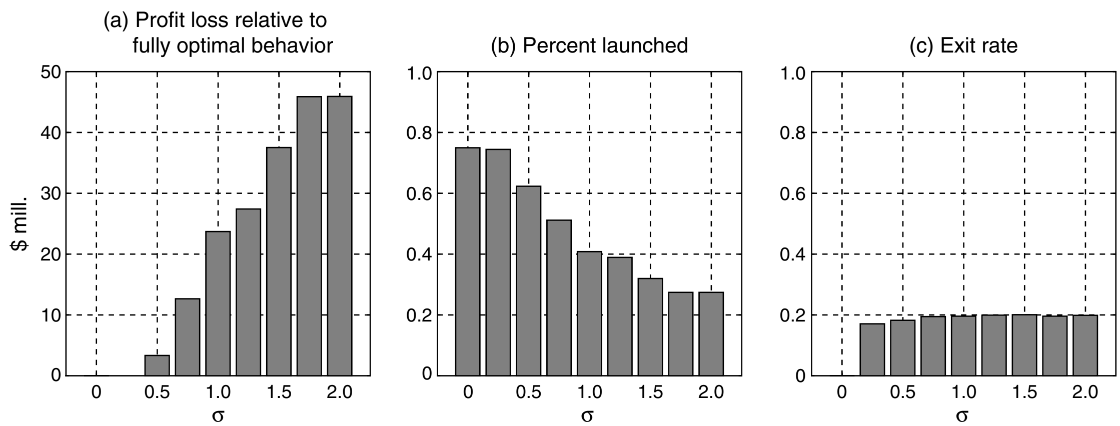
Suboptimal Behavior: Avoiding "Product Failures." In our final counterfactual experiment, we compare optimal product launch behavior to an alternative strategy, where a manager launches a product only if the "failure" probability is below a threshold, or, equivalently, tries to keep the average product exit rate below that threshold. If a large exit rate is taken as indicative of "bad" decision making, this alternative decision rule might be optimal from the manager's point of view. In particular, a manager whose career prospects are influenced by the number of product launches that resulted in a failure will optimally avoid a high exit probability. However, we already know that large exit rates can be consistent with fully optimal behavior, where optimality is measured with respect to the expected PDV of profits as the objective function. Limiting product exit is analogous to a "frequentist" approach to testing if the product is profitable, with an upper bound on the Type II error probability imposed.

Figure 9 shows the profit loss, entry, and exit rates under the alternative product launch procedure. In order to avoid exceeding an exit threshold of 20%, the firm decreases the rate at which new products are launched relative to fully optimal behavior. The resulting profit loss can be huge. For example, at $\sigma_0 = 1.31$, the loss is \$27.5 million; at $\sigma = 2$, it is \$45.9 million. We conclude that if a firm's or a manager's product development and introduction process is judged by the "failure rate" of new products, the resulting change in managerial decision making can strongly decrease profits.

6. Conclusions

This paper deals with the classic marketing problem of optimal new product launch. We specifically focus on product launch under demand uncertainty and its consequences for the optimal launch and product exit decision, and the price a firm should pay to reduce its demand uncertainty by a given amount. Formally, a product launch under demand uncertainty is akin to a sequential experimentation problem in statistical decision theory, where the firm uses the market as a laboratory to acquire information about the true demand for its product from observed sales.

Figure 9 Suboptimal Decision Making: Products Are Not Launched If the Expected Exit Rate Exceeds 20%



Note. Products are drawn from the empirical distribution of all products in the sample.

We develop a model that incorporates this sequential experimentation aspect and predicts the optimal path of stay-or-exit decisions, advertising, and (equilibrium) prices. The optimality of these decisions is judged with respect to the firm's generic objective function, the expected PDV of profits, which is derived from an underlying random coefficients logit demand system. The model is solved using numerical dynamic programming techniques. The model can be solved on a modern personal computer, and the solution can be used as a decision tool to aid managerial decision making during an actual product launch. The model parameters are estimated using a technique that accounts for an econometric endogeneity problem in advertising, and furthermore allows us to recover the firms' initial degree of demand uncertainty. A firm that uses the model as a decision tool faces a simpler estimation task and only needs to estimate the demand parameters—for example, using the BLP method.

Our model is illustrated for the case of the U.S. RTE breakfast cereal industry. We find that in this specific example, the value of reducing demand uncertainty can be substantial. An important strategic lesson is gained by examining how the launch and exit decisions depend on the current degree of demand uncertainty. We discussed the crucial distinction between the loss function and the value function. Under higher demand uncertainty, the average profit loss due to "mistaken" launch-or-exit decisions increases, but the value function of the product also increases. Hence, a decision maker who confuses the value function with the loss function will launch fewer (scrap more) products under higher uncertainty. The opposite behavior, however, is optimal.

A stylized fact of the cereal industry is that a large fraction of all newly launched products "fail," i.e., exit soon after product launch. The optimality of an observed failure rate can be judged using our model. For the example of the RTE cereal industry, we predict that firms should strongly increase the fraction of new product opportunities launched for even small degrees of uncertainty. Hence, the optimal test of the null hypothesis, $H_0 =$ "the product is profitable," largely avoids Type I errors, i.e., scrapping or not launching a truly profitable product. Consequently, firms should optimally tolerate high product exit rates (Type II errors). In fact, we find that the observed product failure rate in the cereal industry is approximately optimal, given the estimated demand uncertainty of the firms in our sample.

If there is a positive fixed cost of entry, the percent of products launched decreases relative to the entry rate under no fixed cost. However, the entry rate still increases strongly in the degree of uncertainty.

We repeatedly stressed that the value from a new product is more than the expected PDV of profits under the current prior. A firm also needs to value the additional information about the product's profitability that it can acquire by delaying exit. This information is valuable to the firm because it has the option of scrapping the product at any time in the future, conditional on its posterior. Our simulations show that the value of additional information is economically important, and can lead to a large profit loss if neglected.

Finally, our results imply that the stakeholders in some firms should not only be concerned about high product exit rates, but also about low exit rates. A manager who, due to career concerns, tries to limit the probability of a product failure may reduce the rate at which new product opportunities are launched to a suboptimal level. This can have a strong negative impact on profits. Hence, a low product "failure" rate per se should not be taken as indicative of optimal managerial decision making.

Beyond the normative use of our model, it also provides a new, alternative explanation for the high observed product entry and exit rates in the U.S. RTE cereal industry. As we already mentioned, the observed launch and exit rates are optimal, conditional on the estimated degree of demand uncertainty. Of course, because we do not observe the cost of reducing the uncertainty, we do not know whether the initial prior uncertainty was chosen optimally. We would like to stress that although we provide an alternative explanation for the stylized facts in the cereal industry, we are unable to test between our explanation and the leading alternative explanation based on product proliferation as entry deterrence. A direct test between the alternative explanations is left for future research.

A main limitation of our model is that we do not explicitly take into account strategic behavior with respect to decision variables that have consequences over time. For example, we do not solve for a dynamic advertising equilibrium, and we do not allow for the possibility that a new product launch may result in competitive product entry or exit. We have argued that in the specific application of our model to the case of the U.S. RTE breakfast cereal industry, the omission of dynamic advertising competition is unlikely to affect the main conclusions. The justification of this claim was based on the particular market structure of this industry, in particular the presence of a large number of competing products. In other markets, however, the explicit consideration of dynamic strategic interaction will be important. Incorporating such strategic aspects into a model of product launch or product entry is left as an important area for future research.

Acknowledgments
Earlier work on this project started with the author's Ph.D. dissertation at Yale University. The author is greatly indebted to his dissertation advisors, John Rust, Steve Berry, and Chris Timmins, for their guidance and support. The author thanks Prof. Ronald Cotterill, the director of the Food Marketing Policy Center at the University of Connecticut, for providing him with data and many insights on the U.S. ready-to-eat breakfast cereal industry. He is greatful to Peter Rossi for his many helpful and insightful comments at various stages of this project. The author has also greatly benefitted from the comments of the area editor and two anonymous reviewers, and discussions with Pat Bayer, Pradeep Chintagunta, Eugene Choo, Tim Conley, J. P. Dubé, Ricky Lam, Puneet Manchanda, Oleg Melnikov, Harikesh Nair, and Peter Reiss. Seminar participants at Berkeley, Columbia, Cornell, MIT, Stanford, the University of Connecticut, the University of Chicago, University of Pennylvania, and Yale provided helpful comments. This research was supported by the Kilts Center of Marketing and a True North Communications Inc. Scholarship at the Graduate School of Business, University of Chicago.

References
Ackerberg, D. A. 2003. Advertising, learning, and consumer choice in experience good markets: A structural empirical examination. *Internat. Econom. Rev.* **44**(3) 1007–1040.

Aghion, P., P. Bolton, C. Harris, B. Jullien. 1991. Optimal learning by experimentation. *Rev. Econom. Stud.* **58** 621–654.

Akçura, M. T., F. F. Gönül, E. Petrova. 2004. Consumer learning and brand valuation: An application on over-the-counter drugs. *Marketing Sci.* **23**(1) 156–169.

Arrow, K. J., M. Nerlove. 1962. Optimal advertising policy under dynamic conditions. *Economica, New Series* **29** 129–142.

Bergemann, D., J. Välimäki. 1997. Market diffusion with two-sided learning. *RAND J. Econom.* **28**(4) 773–795.

Berry, S. T. 1992. Estimation of a model of entry in the airline industry. *Econometrica* **60**(4) 889–917.

Berry, S. T. 1994. Estimating discrete-choice models of product differentiation. *RAND J. Econom.* **25**(2) 242–262.

Berry, S. T., J. Levinsohn, A. Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* **63**(4) 841–890.

Bresnahan, T. F., P. C. Reiss. 1991. Entry and competition in concentrated markets. *J. Political Econom.* **99**(5) 977–1009.

Ching, A. T. 2002. Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration. University of Toronto, Ontario, Canada.

Churchill Jr., G. A., 1995. *Marketing Research*, 6th ed. The Dryden Press, Chicago, IL.

Cotterill, R. W. 1999. High cereal prices and the prospects for relief by expansion of private label and antitrust enforcement. *Agribusiness* **15**(2) 229–245.

DeGroot, M. H. 1969. *Optimal Statistical Decisions*. McGraw-Hill, New York.

Dixit, A. 1989. Entry and exit decisions under uncertainty. *J. Political Econom.* **97**(3) 620–638.

Dixit, A. K., R. S. Pindyck. 1994. *Investment Under Uncertainty*. Princeton University Press, Princeton, NJ.

Dorfman, R., P. O. Steiner. 1954. Optimal advertising and optimal quality. *Amer. Econom. Rev.* **44** 826–836.

Erdem, T., M. P. Keane. 1996. Decision-making under uncertainty: Capturing brand choice processes in turbulent consumer goods markets. *Marketing Sci.* **15**(1) 1–20.

Erdem, T., S. Imai, M. Keane. 2003. Brand and quantity choice dynamics under price uncertainty. *Quant. Marketing Econom.* **1** 5–64.

Gönül, F., M. Z. Shi. 1998. Optimal mailing of catalogs: A new methodology using estimable structural dynamic programming models. *Management Sci.* **44**(9) 1249–1262.

Gönül, F., K. Srinivasan. 1996. Estimating the impact of consumer expectations of coupons on purchase behavior: A dynamic structural model. *Marketing Sci.* **15**(3) 262–279.

Harrington, J. E. 1995. Experimentation and learning in a differentiated-products duopoly. *J. Econom. Theory* **66** 275–288.

Hausman, J. A. 1996. Valuation of new goods under perfect and imperfect competition. T. F. Bresnahan, R. J. Gordon, eds. *The Economics of New Goods*. University of Chicago Press, Chicago, IL, 209–237.

Hendel, I., A. Nevo. 2005. Measuring the implications of sales and consumer stockpiling behavior. NBER Working Paper 11307.

Judd, K. L. 1985. Credible spatial preemption. *RAND J. Econom.* **16**(2) 153–166.

Judd, K. L., M. H. Riordan. 1994. Price and quality in a new product monopoly. *Rev. Econom. Stud.* **61**(4) 773–789.

Kadiyali, V. 1996. Entry, its deterrence, and its accommodation: A study of the U.S. photographic film industry. *RAND J. Econom.* **27**(3) 452–478.

Lehmann, D. R., S. Gupta, J. H. Steckel. 1998. *Marketing Research*. Addison-Wesley, Reading, MA.

Little, J. D. C. 1966. A model of adaptive control of promotional spending. *Oper. Res.* **14**(6) 1075–1097.

McLennan, A. 1984. Price dispersion and incomplete learning in the long run. *J. Econom. Dynam. Control* **7** 331–347.

Melnikov, O. 2001. Demand for differentiated durable products: The case of the U.S. computer printer market. Manuscript, Yale University, New Haven, CT.

Montgomery, A. L., E. T. Bradlow. 1999. Why analyst overconfidence about the functional form of demand models can lead to overpricing. *Marketing Sci.* **18**(4) 569–583.

Moscarini, G., L. Smith. 2001. The optimal level of experimentation. *Econometrica* **69**(6) 1629–1644.

Nevo, A. 2000. A practitioner's guide to estimation of random-coefficients logit models of demand. *J. Econom. Management Strategy* **9**(4) 513–548.

Nevo, A. 2001. Measuring market power in the ready-to-eat cereal industry. *Econometrica* **69**(2) 307–342.

Pekelman, D., E. Tse. 1980. Experimentation and budgeting in advertising: An adaptive control approach. *Oper. Res.* **28**(2) 321–347.

Raman, K., R. Chatterjee. 1995. Optimal monopolist pricing under demand uncertainty in dynamic markets. *Management Sci.* **41**(1) 144–162.

Rao, V. R., F. W. Winter. 1981. A Bayesian approach to test market selection. *Management Sci.* **27**(12) 1351–1369.

Roberts, J. H., G. L. Urban. 1988. Modeling multiattribute utility, risk, and belief dynamics for new consumer durable brand choice. *Management Sci.* **34**(2) 167–185.

Roberts, K., M. L. Weitzman. 1981. Funding criteria for research, development, and exploration projects. *Econometrica* **49**(5) 1261–1288.

Rust, J. 1987. Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica* **55**(5) 999–1033.

Schmalensee, R. 1978. Entry deterrence in the ready-to-eat breakfast cereal industry. *Bell J. Econom.* **9** 305–327.

Song, I., P. K. Chintagunta. 2003. A micromodel of new product adoption with heterogeneous and forward-looking consumers: Application to the digital camera category. *Quant. Marketing Econom.* **1**(3) 371–407.

Trefler, D. 1993. The ignorant monopolist: Optimal learning with endogenous information. *Internat. Econom. Rev.* **34**(3) 565–581.

Urban, G. L., J. R. Hauser. 1993. *Design and Marketing of New Products*, 2nd ed. Prentice Hall, Englewood Cliffs, NJ.

Urban, G. L., G. M. Katz. 1983. Pre-test-market models: Validation and managerial implications. *J. Marketing Res.* **20**(3) 221–234.

Wald, A. 1945. Sequential method of sampling for deciding between two courses of action. *J. Amer. Statist. Association* **40** 277–306.