Consumer Learning about Experience Goods: Evidence from an Online Grocer[*]

Karen Clay[†], Ronald Goettler[‡], and Eric Wolff[§]

April 11, 2003

Preliminary

Abstract

Consumers often have incomplete information when making purchases. For "experience" goods, uncertainty can only be resolved through consumption. If the experience signal is noisy, repeat purchases may be needed to gain precise information. Nelson (1970) models consumer behavior in this setting and establishes the role for experimentation by consumers and the impact on firm and industry behavior. In this paper, we study the impact of incomplete information on consumer behavior when facing a monopolist's menu of two-part tariffs. Using household data from an online grocer of tariff and usage choices over 70 weeks, we estimate a dynamic structural model with forward-looking consumers. We use the model to assess various theories of consumer learning, the impact on welfare due to uncertainty, and the firm's use of two-part pricing menus to segment consumers and extract surplus.

[*]We gratefully acknowledge insightful comments from Shane Greenstein, Michael Katz, Bob Miller, Dan Rabinowitz, Warren Whatley, and Laurie Zinick, and excellent research assistance by Roger Reed and Stanley Ouyang. We also thank the online grocer for providing the data.

[†]Heinz School of Public Policy and Management, Carnegie Mellon University, karen.clay@cmu.edu.

[‡]Graduate School of Industrial Administration, Carnegie Mellon University, ronald.goettler@cmu.edu.

[§]Graduate School of Industrial Administration, Carnegie Mellon University, eric.wolff@cmu.edu.

1 Introduction

Consumers often have incomplete information when choosing which goods and services to purchase. In some cases, consumers can search for the information, but in the case of "experience" goods the uncertainty can only be resolved through consumption of the good. Nelson (1970) models consumer behavior in this setting and establishes the role for experimentation by consumers and the impact on firm and industry behavior.

Recent theoretical models have explored the relationship between ex post learning from one's own experiences versus learning from ex ante sources, such as others' experiences or advertising (McFadden and Train 1996 and Bolton and Harris 1999). There has, however, been comparatively little empirical work on learning. In the marketing literature, Erdem and Keene (1996) investigate the role of advertising in consumer learning about goods. Ackerberg (2002) also studies advertising in a dynamic context, focusing on the distinction between informative advertising and prestige or image advertising. Crawford and Shum (2002) estimate how rapidly consumers learn about the effectiveness of anti-ulcer drugs based on their own experiences.

In this paper, we investigate ex ante and ex post learning by consumers who enrolled with a monopoly online grocer and faced a menu of two-part tariffs for delivery. Using data from the online grocer, we estimate a dynamic structural model with forward-looking consumers. Our model builds on the dynamic brand choice model of Eckstein, Horsky, and Raban (1989), which introduced Bayesian learning to consumer choice in a fully dynamic context. Our model extends theirs in three dimensions: i) consumers face a menu of two-part tariffs, ii) quitting can be either endogenous (due to poor match) or exogenous (due to a move), and iii) consumers can belong to one of three latent classes.

The data include all 6,220 households that signed up for the service in a Midwestern market between September 1997 to January 1999. When the online grocer entered this market, it partnered with one of the two major local supermarket chains. Product prices online were the same as those charged by the partner chain. When consumers signed up for the service, they selected one of three ex ante two-part tariffs with different monthly fees and per delivery charges. Consumers could change plans at any time.

We have three main findings. First, consumers had heterogeneous match values, beliefs, learning rates, and price sensitivities. Furthermore, this heterogeneity exhibits low correlation with demographic data. Second, consumers were generally price insensitive, suggesting that the monopolist could have earned more revenue had it charged higher prices. Simulations show that indeed the monopolist could have increased revenues by omitting the plan with highest monthly fixed fee and no delivery charge. Finally, in line with the predictions of the literature on ex ante learning from others or from advertising, later adopters had lower variance and higher means than did earlier adopters.

Section 2 presents our model of consumer learning. Section 3 discusses econometric issues. Section 4 presents the data and discusses how consumers actually use this service. Section 5 presents estimates of the model, and section 6 investigates price discrimination via a set of policy experiments.

2 Model

We model the household's decision of whether to use the online grocer or traditional grocers.[1] One could imagine modeling this decision on each shopping occasion, as well as the endogeneity of shopping frequency. This is not possible given our data, since we do not observe the use of traditional grocers. Instead, we assume households buy groceries at least once per week, and we model whether they use the online grocer on at least one of these occasions. Throughout the paper, we'll speak as if households purchase groceries exactly once per week, from either the online grocer or a traditional grocer.

There are two dynamic aspects of this decision. First, the online grocer is a new service about which consumers have limited information. We model online grocery delivery as an "experience" good. As the consumer uses the good, she learns, in a Bayesian fashion, whether it is a good match for her. If her prior belief suggests the product is not good, she may still try it since the lower expected current utility from the online grocer may be offset by the possibility of learning that it is in fact good.

The second dynamic aspect arises from the online grocer's use of subscription plans. Consumers choose whether to pay a high delivery charge with no monthly fee, a moderate delivery charge with a moderate monthly fee, or no delivery charge with a high monthly fee. The vector of delivery prices is $p = (11.95, 5, 0)'$ and fees converted to a weekly basis are $F = (0, 6.95, 24.95)'/4.33$. Each consumer chooses the best plan given her beliefs about future usage. If consumers could costlessly switch plans, then this feature of the market would be transparent. However, switching is not costless, since it requires thinking about which plan is best and calling the online grocer to request the change. The switching costs contribute to the dynamic nature of the decision.

Each week the consumer chooses her online grocer usage and subscription choice to maximize the expected discount flow of utility from grocery consumption, net of switching costs. Her expected utility flow from using the online grocer in a given period is

$$u_{online} = \mu - \alpha p' S + \epsilon_{online} \; , \tag{1}$$

where μ is her current mean belief of the online grocer's quality (for her), S is a dummy

[1] Since we only observe households that have subscribed to Peapod, our analysis is conditional on this choice.

vector with 1 in the location of her current subscription plan, and ϵ_{online} represents components that are unobservable to the econometrician, though observable to the consumer prior to each period's usage choice. Utility from using the traditional grocer is simply

$$u_{Store} = \epsilon_{Store} \, . \tag{2}$$

Without subscriptions, we could proceed directly to the Bellman equation to express the consumer's dynamic optimization recursively. In that simpler setting we could write

$$V(\mu) = \max \left\{ u_{online} + \beta \int V(\mu') P(d\mu'|\mu), \; u_{Store} + \beta V(\mu) \right\}, \tag{3}$$

where (in a slight abuse of notation) μ denotes all the belief parameters, and $P(\mu'|\mu)$ is the distribution of next period's beliefs conditional on current beliefs. To account for subscription choice, we add the state variable $S \in \{(1,0,0)', (0,1,0)', (0,0,1)'\}$. The Bellman equation then becomes

$$V(\mu, S) = \max \left\{ \begin{array}{l} u_{online} + \beta \int \max\limits_{S'} \left\{ V(\mu', S') - \alpha F' S' - \delta I(S \neq S') \right\} P(d\mu'|\mu), \\ u_{Store} + \beta V(\mu, S) \end{array} \right\} \tag{4}$$

where $I()$ is the indicator function and δ denotes switching costs. Note that the optimal subscription is chosen after the experience signal is observed and beliefs are updated. Conceptually, the consumer may change S after using the traditional grocer. However, in the above model she would never do so. The optimal S given current beliefs would already have been chosen immediately after her previous the online grocer usage, since her beliefs do not change unless the online grocer is experienced.[2]

While subscription choice is deterministic given beliefs, this choice is probabilistic for the econometrician who does not observe beliefs. Our results thus far suggest that this "flexibility" is too limited. That is, the likelihood function is ill-behaved unless we add an additional stochastic component to the consumer's subscription choice.[3] For simplicity, we add an unobservable type I extreme value utility to the conditional continuation values for each S' in the maximization over S' in (4). Since the online grocer usage is chosen prior to observing these shocks, the consumer must integrate over their possible values to determine the continuation value. Randomness in the S' choice implies that changing plans may be desired immediately after using a traditional grocer. Using the property that the maximum of several extreme value random variables

[2] While one could model usage and subscription as simultaneous choices, we feel most consumers change their subscriptions in response to recent experiences, not in response to a current draw from $_{online}$.

[3] Jumping ahead to the estimation discussed in the next section, the problem is that, conditional on some vector of parameters, for some consumers none of the Monte Carlo draws of initial beliefs leads to the initially chosen plan as being optimal.

also has an extreme value distribution, the Bellman equation becomes

$$V(\mu, S) = \max \left\{ \begin{array}{l} u_{online} + \beta \int (\gamma + \exp(\sum_{S'} \ln(V(\mu', S') - \alpha F'S' - \delta I(S \neq S'))))P(d\mu'|\mu), \\ u_{Store} + \beta(\gamma + \exp(\sum_{S'} \ln(V(\mu, S') - \alpha F'S' - \delta I(S \neq S')))) \end{array} \right\} \tag{5}$$

where γ is Euler's constant (the mean of a type 1 extreme value random variable).

The Markov transition kernel $P(\mu'|\mu)$ is determined by the specification of initial beliefs and learning. We specify the consumer's prior belief of the online grocer's "quality" or "convenience" to be $N(\mu_0, \sigma_0^2)$. The (noisy) experience signal, μ_e on each peapod usage is assumed to be distributed $N(\bar{\mu}, \sigma_e^2)$. Since the normal distribution is a conjugate prior for normal random samples, the Bayesian updating of beliefs has a simple closed form. Assuming known σ_e^2, current prior mean μ after n experiences, and (new) experience signal μ_e, the posterior mean is

$$\mu' = \frac{\sigma_e^2 \mu + \sigma^2(n)\mu_e}{\sigma_e^2 + \sigma^2(n)}, \tag{6}$$

where

$$\sigma^2(n) = \frac{\sigma_e^2 \sigma_0^2}{\sigma_e^2 + n\sigma_0^2} \tag{7}$$

is the deterministic evolution of the variance of beliefs.

V in (5) is a fixed point obtained by iterating the corresponding contraction mapping. Given beliefs are normal, the integral over μ' is efficiently evaluated using Gauss-Hermite quadrature. Beliefs are discretized and linear interpolation is used to evaluate V at arbitrary beliefs.[4]

3 Estimation

We assume ϵ_{online} and ϵ_{Store} are i.i.d type I extreme value. As explained by Rust (1987), this assumption satisfies the *conditional independence* needed for maximum likelihood estimation of the structural parameters. We first present the likelihood of a household's choices over T periods conditional on the unobserved beliefs. Then we specify the distribution of beliefs across the population and present the likelihood, integrating over beliefs. The likelihood, conditional on beliefs, is given by

$$L(S_1, \ldots, S_T, d_1, \ldots, d_T | \mu_1, \ldots, \mu_T, \theta, S_0) = \prod_{t=1}^{T} P_S(S_t|S_{t-1}, \mu_t) P_d(d_t|\mu_t, S_t), \tag{8}$$

where P_S is the conditional choice probability of the subscription plan, P_d is the conditional choice probability of the usage choice, $d \in \{online,\ Store\}$, and θ is the parameter

[4] Convergence requires about one minute on a 1.1 gigahurtz AMD Athlon, starting from $V = \max\{u_{online}, u_{Store}\}$ when $\beta = 0.995$ and other parameters have sensible values.

vector (supressed on the right-hand side for compactness). Let $S_0 = 0$ denote the consumer's lack of association with any subscription plan upon starting the process. Both P_S and P_d have the familiar logit formula

$$P_S(S_t|\mu_t, S_{t-1}) = \frac{\exp(V(\mu_t, S_{t-1}) - \alpha F' S_t' - \delta I(S_t \neq S_{t-1}))}{\sum_S \exp(V(\mu_t, S_{t-1}) - \alpha F' S' - \delta I(S \neq S_{t-1}))} \tag{9}$$

and

$$P_d(d_t|\mu_t, S_t) = \frac{\exp(u_{d_t} + \beta V_{d_t}(\mu_t, S_t))}{\sum_d \exp(u_d + \beta V_d(\mu_t, S_t)))}, \tag{10}$$

where V_d is the continuation value. For $d = online$ the continuation value is the first line of (5) and for $d = Store$ it is the second line of (5).

3.1 Exogenous quits and censored plans

Many households in the data use the online grocer regularly over a long period of time, and then suddenly stop. While such behavior could be explained by a slow learning process, it more likely reflects permanent household shocks, such as moving from the area, marriage, divorce, childbirth, retirement, etc. To account for such shocks, we assume that each household in each period ***exogenously quits*** with probability γ.

We model consumers as being unaware of the possibility of these shocks. That is, we do not modify the dynamic choice problem to account for a possible truncation of the infinite horizon. This is consistent with consumer's having a (biased) prior of $\gamma = 0$, or with expecting an online grocer in their new location (in the event of moving).

The conditional likelihood in (8), augmented with an exogenous quit rate γ, becomes

$$L(\cdot) = \left[\prod_{t=1}^{T_q} (1-\gamma) P_S(S_t|S_{t-1}, \mu_t) P_d(d_t|\mu_t, S_t)\right] Prob(T - T_q \text{ trailing periods of no usage}) \tag{11}$$

where T_q is the last period of online usage, and $L(\cdot)$ has the same arguments as in (8). The last factor in (11) is the probability of observing $T - T_q$ periods of no usage at the end of a household's recorded data.

Each period's inactivity may be due to either quitting in the current period or a prior period, or due to choosing the traditional store despite being a the online grocer subscriber. Once the consumer quits, however, the subsequent periods of inactivity no longer represent observations. Hence, if $T = 52$ and $T_q = 50$, the probability of the last two periods is the sum of the probability she exogenously quits at $t = 51$, the probability she chooses no usage at $t = 51$ and exogenously quits at $t = 52$, and the probability she chooses no usage at both $t = 51$ and $t = 52$.

The consumer's subscription plan is also censored for these trailing weeks since we only observe the plan when an order is processed. Hence, computing the probability

of no usage requires integrating over the unobserved plans. For households with long trailing periods, this integration becomes an exploding tree. To avoid this, we (arbitrarily) assume that if $T - T_q > 4$ the consumer chooses the plan with no monthly fee in period $T_q + 1$ and stays on it thereafter. If $T - T_q <= 4$ the consumer remains on her previous subscription plan for the remaining periods.[5]

Accounting for exogenous quits and censored trailing periods, the probability of the trailing period of no usage is

$$\sum_{t=T_q}^{T} \gamma^{I(t<T)}((1-\gamma)P_d(d_t|\mu_t, S_t^c))^{t-T_q} P_S(S_{T_q+1}^c|S_{T_q}, \mu_{T_q+1})^{I(t>T_q)} P_S(S_t^c|S_{t-1}^c, \mu_t))^{\max(0,t-T_q-1)}, \tag{12}$$

where the superscript c indicates censored subscriptions. This expression is simply the sum of the probabilities of the many ways in which an exogenous quit can be combined with periods of no usage despite not having yet quit. Using the above example with two trailing periods this sum is $\gamma + \gamma(1-\gamma)P_d P_S + (1-\gamma)^2 P_d^2 P_S^2$, where the P_d and P_S are the same in both periods since the subscription and beliefs are unchanged. In the second line, the first P_S is for the probability of the first trailing period, which *may* involve a change in S, and the second P_S is for the remaining periods, which have no changes in S (by assumption).

3.2 Beliefs

Beliefs entail three unobservable components: the household's true online grocer quality $\bar{\mu}$, the household's prior μ_0, and the vector of noisy experience draws μ_e. We estimate the distribution of these parameters, assuming they are i.i.d. normal across the population of households. Adding a subscript h to denote household specific values to the belief notation used above,

$$\begin{aligned} \bar{\mu}_h &\sim N(\bar{\mu}_H, \bar{\sigma}_H^2) \\ \mu_{0_h} &\sim N(\bar{\mu}_h + \mu_{0_H}, \sigma_{0_H}^2) \\ \mu_{e_h} &\sim N(\bar{\mu}_h, \sigma_{e_H}^2), \end{aligned} \tag{13}$$

where the subscript H denotes values common across households. Note that experience signals are unbiased, given each household's $\bar{\mu}_h$, though the prior mean is biased if $\mu_{0_H} \neq 0$. Using R monte carlo draws from the distribution in (13) to numerically

[5] If the plan choice were deterministic given beliefs, as specified in (4), then any subscription change would occur in the first trailing week and the integration would only have three mass points in every trailing week. Currently, we are trying to estimate the model without the random shock in subscription choice.

integrate over these unobservables, the simulated likelihood for a given household is

$$\tilde{L}(S_1,\ldots,S_T,d_1,\ldots,d_T|\theta,\{\bar{\mu}^r,\mu_0^r,\mu_e^r\}_{r=0}^R)=\frac{1}{R}\sum_{r=1}^{R}L(S_1,\ldots,S_T,d_1,\ldots,d_T|\mu_0^r,\ldots,\mu_T^r,\theta) \tag{14}$$

where beliefs evolve according to (6) and (7) using the simulated μ_e. Our estimator is obtained by maximizing the product of the households' simulated likelihoods, using the nested fixed-point algorithm of Rust (1987).

3.3 Household heterogeneity

Thus far we have avoided using observable household characteristics. Allowing β, α, δ, or $\sigma_{e_H}^2$ to vary across households requires finding the fixed point V for each set of these parameters. For example, using two demographic variables, each with say three values, increases computation time by approximately a factor of nine, for each evaluation of the likelihood function. Given the number of household characteristics that we are interested in analyzing, simply interacting them with model parameters is not computationally feasible.

Instead, we specify a limited number of latent classes, across which all model parameters vary. After estimation of θ, we compute each household's posterior probability of belonging to each latent class and regress these probabilities on the observed characteristics.[6] This latent class approach suits our study well, since we find that much of the heterogeneity across households is not explained by observable characteristics.

4 Data

We have data on all 6,220 households that enrolled with the online grocer during the 16-month period from September 16, 1997 to January 23, 1999. Customers learned about the service through advertising, in the form of mass mailings, media stories, print and radio advertising, in store advertising in the partner chain, and displays on the delivery trucks. Customers almost certainly also learned about the existence and value of the service from conversations with others or from newspaper reports.

More than half the users voluntarily provided demographic information at the time they signed up for the service.[7] This demographic information included household structure, age of the subscriber, and income. Comparing households with and without demographic data, we find the two groups do not differ significantly on dimensions such as enrollment date, plan choice, and usage. Thus, the households that provided

[6] One could also enable the prior probabilities, used to integrate over the latent classes in the estimation of θ, to depend on these characteristics.

[7] Unfortunately, no question asked how the customer heard about the service.

demographic data appear to be representative of the larger population of customers that used the service. Once they had enrolled, customers placed orders on a computer using software or a web-based interface. The monopolist offered online grocery service in conjunction with an existing local grocery chain. Prices for the groceries were the same as the prices in the customer's local store. Customers selected a two-hour delivery window, typically the next day, when someone could be home to accept the delivery.

4.1 Adopters vs. the General Population

Table 1 compares the demographic characteristics of adopters and the demographic characteristics of the county that the grocer served.[8] Compared to the county overall, adopting households had more adults, more children, and higher incomes. The median household income in the county was almost $40,000 per year. Of adopters reporting incomes, only 30 percent were in the lowest income group (under $50,000). Adopters were also more likely to be in the 35–50 age range and less likely to be in the over 50 age range.

These demographic differences are not surprising for a number of reasons. First, consumers had to have computers and modems that they could use to place the order. Census results for October 1997 indicate that Internet usage ranged from less than 10 percent for households with income of $15,000 to less to more than 40 percent for households with income of $75,000 or more. Thus, we would expect to see demographic differences between adopters and nonadopters. Second, the differences are consistent with stylized facts from the diffusion literature that adopters usually have higher income and larger households. Third, the online grocery industry explicitly targets busy, dual-income households, so we would expect these types of households to be a large fraction of the customers.

The adoption pattern over time follows the classic s-shaped curve. Figure 1 shows the cumulative enrollments over our data period. The s-shape may reflect a number of different factors including social learning, internet penetration, and the intensity of advertising over time.

4.2 Subscription Plan Choice

The monopolist offered customers a choice of three two-part tariffs.[9] Table 2 describes the fixed and marginal delivery costs that customers faced under the three plans. The prices in these plans were constant throughout the period. The lower envelope of the

[8] All comparisons are based Census 2000 numbers for the county that the online grocer served.

[9] Most online grocers o er a single fee structure, choosing either delivery fees or monthly charges. Some o er delivery charges declining in the size of the order.

three plans was convex and so was ex ante incentive compatible. That is, lower demand customers would not prefer to choose plans designed for high demand customers and vice versa. The relatively high delivery fee for the low plan (no monthly charge) suggests that the monopolist did not plan to serve a large part of the population. Penetration rates overall were 0–3 percent of households for most zip codes, although penetration was a higher fraction of those who had Internet access. Census survey results indicate that only 22 percent (33 percent) of the population had internet access from any source in October 1997 (December 1998).[10] The most significant constraint on pricing was that customers would simply go to the grocery store themselves.

Table 3 shows the demographic characteristics of customers across plans. Households on the frequent and moderate usage plans appear to have been larger, to have had more children, and to have been higher income than households on the low usage plan. Both early and late subscribers were also more likely to sign up for the frequent and moderate usage plans than are subscribers who sign up during months 4–12.[11]

4.3 Usage

As customers used the service, they learned about its value for them. The learning was likely to be substantial, since online grocery shopping differs from in-store shopping in a number of respects. A typical in-store shopper shops without a list, spends $35 per trip to the grocery store, and shops two to three times per week. In contrast, online customers had to make a shopping list. The first time they shopped online, they also had to know the specific names of desired items since the traditional cues, such as shelf position and package size, are absent. To simplify subsequent orders, the online grocer maintained records of prior orders. Delivery also required adjustment. Customers had to wait at least one day between order and delivery, and an adult had to be home to accept delivery of the groceries. Clearly, the value of online shopping depended on how difficult the household found these order and delivery requirements.

Table 4 presents summary statistics on usage. Many people signed up but never used the service. The vast majority of these customers had signed up for the low usage plan, so their only loss was the time spent filling out the form. Retention of active users on all plans was poor. Within three months, between one-third and one-half of customers on each plan had quit. Customers on plans that were optimal for higher expected purchases did purchase more on those plans. Specifically, users on the frequent plan purchased 3.2 times per month; users on the moderate plan purchased 1.7 times per month, and users on the low plan purchased 1.3 times per month. Average order size

[10] A Nation Online (2002) http://www.ntia.doc.gov/ntiahome/dn/anationonline2.pdf
[11] The reasons for this will be explored in a later section.

of \$114 was substantial relative to in store purchases and about average for the online grocery industry.

Two things are striking about usage: i) many customers discontinued usage after just a few experiences, and ii) ex post mistakes (not being on the cheapest plan) were significantly more common on the frequent plan than on the other two plans. The number of mistakes per user per month was 0.58 on the frequent plan as compared to 0.17 for the moderate plan and 0.30 for the low plan.[12] This is consistent with the finding in telecommunication that flat rate users (equivalent to the frequent plan) make more mistakes than measured rate users (equivalent to the other two plans).[13] This stylized fact has been attributed to flat rate users not wanting to have to monitor usage or to risk aversion.

5 Results

We estimate the model using one, two, and three latent classes. The integration over unobserved beliefs is performed using $R = 100$ draws (per household). The exogonous quit rate, γ is fixed at 0.003 based on census data of household mobility in this county. The discount factor, β is fixed at 0.984.[14]

The standard information criteria (BIC, AIC, CAIC) all suggest the third class is warranted. (We have yet to check whether a fourth is also warranted.) Table 5 reports estimates for the three segment model.

All three segments have upwardly biased prior beliefs, revealed by the positive and significant μ_{0_H} estimates. Furthermore, all consumers learn slowly since σ_{e_H} is high for each segment. The learning rates are further detailed, via the evolution of posterior means and variances as usage increases, in table 6.

The α estimates indicate that consumers are not price sensitive, particularly segments 2 and 3. This general lack of price sensitivity is not surprising given that most of these households have significant disposable income, and are experimenting with the online grocer with the intent of simplifying their lives. These estimates of α suggest the online grocer could have charged higher prices and fees without losing many customers. We explore this possibility in Section 6.

Finally, the δ estimates are all high, reflecting the fact that households rarely switch

[12] The monopolist benefited financially from users' mistakes. The expected cost of mistakes per customer per month was \$4.37 for customers on the frequent usage plan, \$0.94 for customers on the moderate usage plan and \$3.44 for customers on the low usage plan.

[13] See, for example, Hobson and Spady (1988), Kling and Ver Ploeg (1988), Kridel, Lehman, and Weisman (1993), MacKie-Mason and Lawson (1993), and Train, McFadden, and Ben Akiva (1987). In contrast, Miravete (2002) found that measured users made more mistakes than flat rate users.

[14] We experimented with estimating β, yielding values between 0.97 and 0.99, but with high standard errors.

plans.[15] Given the low marginal utility of money (revealed by α), these switching costs are extremely high in dollar terms. For example, segment 1's $\delta = 5.454$ and $\alpha = .092$ suggests switching costs for the most price sensitive segment are equivalent to $5.454/.092 = \$59.28$.

5.1 Household demographics

As discussed in the previous section, we do not use household demographic data in the estimation. However, we can still assess whether behavior differs across demographic groups. Table 7 presents means of various characteristics for the three latent segments. Each household is assigned to the segment for which its posterior probability of membership is highest. The table includes several classes of demographic measures, as well as plan and usage choices. Observable household characteristics are similar across segments. In unreported regressions, we regress the demographics on posterior segment probabilities and find R-squared's ranging from .026 to .053.[16]

In short, households differ but the differences are unrelated to observed characteristics. Given this lack of correlation and our interest in price discrimination, it is important that we have allowed for unobserved heterogeneity.[17]

5.2 Social Learning

Theoretical models of consumer learning, such as McFadden and Train (1996), distinguish between learning from one's own experiences versus "social" learning from others' experiences. In the presence of social learning, some consumers delay experimentation until they learn from others that the product is likely to be a good match for them. While we do not explicitly model strategic delays due to social learning, we can assess whether late adopters had prior beliefs with higher mean and lower variance as predicted by models with social learning.

Table 8 presents estimates of the model (without latent classes) for two sub-samples of households, as well as for the full sample. The "early" sub-sample contains the 1042 households that enrolled in the first 15 (of 70) weeks. The "late" sub-sample contains the 744 households that enrolled during the last 24 weeks. Relative to early adopters, late adopters had higher match values $\bar{\mu}_H$, less dispersion in match values $\bar{\sigma}_H$, and lower prior variance σ_{0_H}. These comparisons are consistent with the social learning

[15] The online grocer claims that customers even ignored letters explaining that they should switch to a different plan.

[16] We also estimate an unreported logit model of initial plan choice on demographics. The pseudo-R-squared is 0.053.

[17] One might also allow this heterogeneity to be drawn from continuous distributions (mixtures of normals) using the importance sampling methodology proposed by Ackerberg (2002b).

hypothesis. We also find that late adopters learn faster from their own experiences (i.e., lower σ_{e_H}), though this is not a prediction of social learning models. The point estimate for the prior's bias, μ_{0_H} is larger for late adopters, though the difference is not statistically significant. Furthermore, the learning rates presented in Table 9 reveal that this bias is quickly corrected by the relatively precise experience signals of late adopters.

6 Policy Experiments

We conducted a number of policy experiments by simulating the model given the estimated parameters from the three latent class model.

One experiment of interest addresses the question: What would have happened had the monopolist only offered a subset of the three two-part tariffs? Recall that the three tariffs were characterized by high ($24.95), middle ($5.00), and low ($0.00) monthly fixed fees. The six experiments were: i) offer the middle and low fee plans, ii) offer the high and low fee plans, iii) offer the high and middle fee plans, iv) offer only the high fee plan, v) offer only the middle fee plan, and vi) offer only the low fee plan. Table 10 shows the total revenue (fees and delivery charges) per household per week for the actual menu of tariffs, and for each of the six experiments.[18]

Removing the middle monthly fee plan generates 12.3 percent more revenue than the actual menu with all three plans. The reasons for this gain will be discussed during the seminar.

References

Ackerberg, D. (2002), "Advertising, Learning, and Consumer Choice in Experience Good Markets: A Structural Empirical Examination," forthcoming International Economic Review.

Ackerberg, D. (2002b), "A new use of importance sampling to reduce computational burden in simulation estimation," manuscript UCLA.

Bergemann, D., and J. Valimaki (1996), "Learning and Strategic Pricing," Econometrica, 64,1125–1150.

[18] These experiments do not address the broader question: What set of two part tari s would have been optimal for the monopolist? We plan to explore this question in future work. In particular, a monopolist may find it best to give provide the service at no charge to ensure that all consumers with high μ_h indeed learn of the service's value. After the learning period, the monopolist would raise prices to maximizize profits. Alternatively, if consumers learn slowly, or have (upwardly) biased priors, then the monopolist's best opportunity to earn profits may be during the learning phase.

Bolton, P., and C. Harris (1999), "Strategic Experimentation," Econometrica, 67, 349–374.

Crawford, G. and M. Shum (2002), "Uncertainty and Learning in Pharmaceutical Demand," Mimeo, University of Toronto.

Eckstein, Z., D. Horsky, and Y. Raban (1988), "An Empirical Dynamic Model of Optimal Brand Choice," Working Paper, Tel-Aviv University.

Erdem, T., and M. Keane (1996), "Decision-making under Uncertainty : Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," Marketing Science, pp. 1–20.

Kihlstrom, R., and M. Riordan (1984), "Advertising as a Signal," Journal of Political Economy, pp. 427–450.

McFadden, D., and K. Train (1996), "Consumers' Evaluation of New Products: Learning from Self and Others," Journal of Political Economy, 104(4), 683–703.

Milgrom, P., and J. Roberts (1986), "Price and advertising signals of product quality," Journal of Political Economy, 94(4), 796–821.

Miller, Robert A. (1984), "Job Matching and Occupational Choice," Journal of Political Economy, 92(6), 1086–1120..

Miravete, E. (2002), "Choosing the Wrong Calling Plan? Ignorance and Learning," American Economic Review, Forthcoming 2002.

Nelson, P. (1970) "Information and Consumer Behavior," Journal of Political Economy, 78, 311–329.

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

Shapiro, C. (1983), "Optimal Pricing of Experience Goods," Bell Journal of Economics, 14(2), 497–507.

Figure 1: Signups over Time

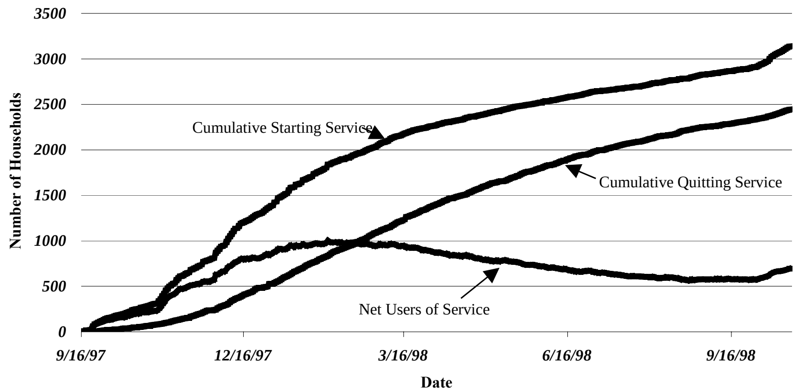

Table 1: Demographic Characteristics of County and Sample

	County	Sample
	Household Structure	
One adult households	44%	14%
Households with children under 18	33%	63%
	Age	
Age, 18-34	38%	40%
Age, 35-50	30%	49%
Age 50+	31%	11%
	Income	
Income	50% < $40,000	30% < $50,000

Notes: One adult households are the sum of householders living alone and female head of household with children, no husband present. This will slightly underestimate the total number one adult households because it excludes male head of household with children. Households under 18 is the unconditional probability. Age distributions were constructed from Census data assuming a uniform distribution of ages across an age span. Income data was a 1997 forecast, because the 2000 income was not yet available.

Table 2: Plan Characteristics

Plan	*Monthly Fee*	*Per Delivery Fee*	*Optimal if usage is*
Frequent	$24.95	$0 for orders > $60, $3.95 otherwise	> 2.9 orders/mo
Intermediate	$5.00	$6.95	1-2.9 orders/mo
Low	$0	$11.95	0-1 orders/mo

Table 3: Demographic Characteristics of Customers by Plan

Characteristics	Overall	Frequent Usage Plan	Moderate Usage Plan	Low Usage Plan
Household Structure				
One Adult, No Children	284	2%	25%	72%
One Adult, Children	190	12%	32%	57%
Avg. Children given 1 adult	0.7	1.7	0.8	0.6
Two Adults, No Children	802	6%	30%	63%
Two Adults, Children	1741	16%	34%	50%
Avg. Children given 2 adults	1.4	1.9	1.5	1.3
>2 Adults, No Children	189	9%	32%	59%
>2 Adults, Children	246	15%	28%	57%
Avg. Children given >2 Adults	1.1	1.3	1.0	1.0
Access to Grocery Stores				
Partner Stores in zip	1.2	1.2	1.2	1.2
Other Chain Stores in zip	1.0	1.0	1.0	1.1
Age of Subscriber				
Age 18-34	1384	11%	33%	57%
Age 35-49	1677	14%	32%	54%
Age 50+	391	6%	30%	63%
Household Income				
0-49,999	808	7%	28%	65%
50,000-69,999	701	9%	27%	64%
70,000-89,999	551	14%	34%	53%
90,000+	647	17%	37%	46%
Not Reported	745	14%	34%	51%
When Signup				
Early, Months 1-3	746	19%	43%	38%
Middle, Months 4-12	2558	10%	28%	63%
Late, Months 13-16	148	17%	45%	38%

Table 4: Usage Characteristics of Customers

	Frequent	Moderate	Low
Active Users	660	1573	1584
Total active months/user	4.6	3.8	2.7
Mean # of Uses/Month	3.2	1.7	1.3
St. Dev., # of Uses/Month	1.5	1.1	1.0
Mean, Purchase Amount	$114.74	$115.00	$120.25
Standard Deviation, Purchase Amount	38.59	54.19	61.36
# Mistakes/User/Month	0.58	0.17	0.30
Mean, Cost of Mistake	$7.54	$5.50	$11.46
St. Dev., Cost of Mistake	5.12	2.46	8.20

Table 5: Parameter Estimates, model with 3 latent classes

Parameter	Segment 1	Segment 2	Segment 3
$\bar{\mu}_H$	-8.249	-3.827	-0.103
	(0.446)	(0.252)	(0.116)
μ_{0_H}	6.836	3.733	0.717
	(0.440)	(0.254)	(0.111)
$\bar{\sigma}_H$	0.175	0.570	0.888
	(0.126)	(0.061)	(0.053)
σ_{0_H}	5.004	1.069	0.890
	(0.241)	(0.053)	(0.055)
σ_{e_H}	6.268	3.161	4.012
	(0.339)	(0.216)	(0.591)
α	0.092	0.010	0.017
	(0.006)	(0.003)	(0.003)
δ	5.454	12.368	15.780
	(0.206)	(2.842)	(3.638)
Segment share	0.351	0.396	0.253
		(0.022)	(0.017)

Asymptotic standard errors in parentheses.

For all segments $\beta = 0.984$ and $\gamma = 0.003$.

Table 6: Posterior Beliefs

cumulative usage	Segment 1 σ	Segment 1 μ	Segment 2 σ	Segment 2 μ	Segment 3 σ	Segment 3 μ
0	5.004	-1.413	1.069	-0.094	0.890	0.614
1	3.911	-4.074	1.013	-0.477	0.869	0.580
2	3.318	-5.244	0.964	-0.789	0.849	0.549
3	2.932	-5.901	0.922	-1.047	0.831	0.521
4	2.656	-6.323	0.885	-1.265	0.813	0.496
5	2.446	-6.616	0.853	-1.452	0.797	0.472
6	2.278	-6.832	0.823	-1.613	0.782	0.450
7	2.141	-6.997	0.797	-1.753	0.768	0.430
8	2.026	-7.128	0.773	-1.877	0.754	0.411
9	1.928	-7.234	0.750	-1.987	0.741	0.394
10	1.843	-7.322	0.730	-2.085	0.729	0.377
11	1.768	-7.396	0.711	-2.173	0.717	0.362
12	1.702	-7.459	0.694	-2.253	0.706	0.347
13	1.642	-7.513	0.678	-2.325	0.695	0.334
14	1.589	-7.560	0.663	-2.391	0.685	0.321
15	1.540	-7.602	0.649	-2.452	0.675	0.309
16	1.495	-7.639	0.635	-2.507	0.666	0.298
17	1.455	-7.672	0.623	-2.559	0.657	0.287
18	1.417	-7.701	0.611	-2.606	0.648	0.277
19	1.382	-7.728	0.600	-2.650	0.640	0.267
20	1.350	-7.752	0.590	-2.691	0.632	0.258
25	1.216	-7.846	0.544	-2.859	0.596	0.218
30	1.116	-7.909	0.508	-2.984	0.566	0.186
35	1.037	-7.956	0.478	-3.080	0.539	0.160
40	0.972	-7.991	0.453	-3.157	0.517	0.138
45	0.919	-8.019	0.431	-3.219	0.496	0.120
50	0.873	-8.041	0.412	-3.271	0.478	0.104
55	0.833	-8.060	0.396	-3.315	0.462	0.090
60	0.799	-8.075	0.381	-3.352	0.448	0.078
65	0.768	-8.088	0.368	-3.384	0.434	0.067
70	0.741	-8.099	0.356	-3.412	0.422	0.058

σ is posterior standard deviation. μ is posterior mean, assuming prior is $\bar{\mu}_H + \mu_{0_H}$ and all experience signals are μ_H.

Table 7: Choices and Characteristics by Latent Class (assignment using highest posterior)

	segment 1	segment 2	segment 3
N households	1110	1178	580
common prior * 100	35.1	39.6	25.3
mean week enrolled	27.9	30.3	25.0
mean week dropped	39.7	58.4	62.5
min week dropped	3.0	6.0	9.0
max week dropped	71.0	71.0	71.0
share order last 5	17.6	58.0	73.8
mean N weeks (nW)	11.9	28.1	37.5
mean N orders	1.9	7.0	22.9
share peapod usage	16.3	24.7	61.1
share No Fee	4.6	15.5	52.7
share Lo Fee	40.5	53.3	38.2
share Hi Fee	54.9	31.2	9.0
share usage — No Fee	50.6	32.4	71.7
share usage — Lo Fee	17.2	23.9	49.6
share usage — Hi Fee	12.7	22.2	48.0
share all demo n.a.	21.4	23.0	17.9
share no demo n.a.	4.1	5.4	4.7
share income n.a.	57.9	58.8	54.1
share income > 90k	22.7	32.6	36.5
share income 50-90k	49.3	44.1	42.1
share income < 50k	28.1	23.3	21.4
mean N adults	2.0	2.0	2.0
mean N kids	1.3	1.4	1.7
share HTML	35.6	40.7	27.7
share v 4	27.5	28.4	37.8
share v 4 HTML	2.3	1.6	1.7
share v 5 SURF	34.6	29.3	32.8
share primary user	91.5	91.0	92.6
share female	69.2	73.6	74.5
share married	77.5	81.4	86.5
share co-habit	6.7	4.6	3.8
share single	15.9	14.0	9.7
share age 18-24	2.2	2.0	1.0
share age 25-44	41.5	36.8	33.2
share age 35-49	46.3	51.9	57.5
share age 50-64	8.4	7.5	6.7
share age 65+	1.6	1.8	1.5
share some HS	0.6	0.8	0.3
share grad HS	9.3	10.3	7.0
share some College	28.8	22.8	20.7
share grad College	42.7	43.5	45.7
share some Grad	18.6	22.6	26.4
share full out	72.9	67.4	67.0
share part out	10.3	12.6	10.9
share full home	12.1	14.0	15.8
share student	1.5	1.3	1.4
share retired/other	3.2	4.7	4.9
share full out, spouse	87.0	88.4	89.4
share part out, spouse	4.4	3.5	2.4
share full home, spouse	2.6	3.5	4.5
share student, spouse	1.4	1.5	0.3
share retired/o, spouse	4.6	3.1	3.3

Table 8: Parameter Estimates, Early versus Late Adopters

Parameter	All Adopters	Early Adopters	Late Adopters
$\bar{\mu}_H$	-1.712	-1.732	-1.293
	(0.039)	(0.048)	(0.165)
μ_{0_H}	1.701	1.390	1.684
	(0.052)	(0.051)	(0.187)
$\bar{\sigma}_H$	1.496	1.751	1.220
	(0.023)	(0.035)	(0.068)
σ_{0_H}	1.842	1.966	0.257
	(0.044)	(0.056)	(0.162)
σ_{e_H}	2.519	2.955	0.175
	(0.068)	(0.110)	(0.100)
α	0.012	0.016	0.008
	(0.001)	(0.002)	(0.003)
δ	12.469	10.351	34.394
	(1.940)	(1.630)	(5.162)

Asymptotic standard errors in parentheses.
For all segments $\beta = 0.984$ and $\gamma = 0.003$.
1042 early adopters enrolling in first 15 weeks.
744 late adopters enrolling in last 24 weeks.

Table 9: Posterior Beliefs: Early versus Late Adopters

cumulative usage	All Adopters σ	All Adopters μ	Early Adopters σ	Early Adopters μ	Late Adopters σ	Late Adopters μ
0	1.842	-0.011	1.966	-0.342	0.257	0.391
1	1.487	-0.604	1.637	-0.768	0.145	-0.758
2	1.280	-0.890	1.432	-0.994	0.112	-0.975
3	1.141	-1.059	1.288	-1.135	0.094	-1.067
4	1.040	-1.171	1.181	-1.230	0.083	-1.117
5	0.961	-1.249	1.097	-1.299	0.075	-1.149
6	0.898	-1.308	1.028	-1.352	0.069	-1.172
7	0.846	-1.354	0.971	-1.393	0.064	-1.188
8	0.802	-1.390	0.922	-1.426	0.060	-1.200
9	0.764	-1.420	0.881	-1.453	0.057	-1.210
10	0.731	-1.444	0.844	-1.476	0.054	-1.218
11	0.702	-1.465	0.811	-1.495	0.052	-1.225
12	0.676	-1.483	0.782	-1.512	0.050	-1.230
13	0.653	-1.499	0.756	-1.526	0.048	-1.235
14	0.632	-1.512	0.733	-1.539	0.046	-1.239
15	0.613	-1.524	0.711	-1.550	0.045	-1.242
16	0.596	-1.534	0.691	-1.560	0.043	-1.245
17	0.580	-1.544	0.673	-1.569	0.042	-1.248
18	0.565	-1.552	0.656	-1.577	0.041	-1.251
19	0.551	-1.560	0.641	-1.584	0.040	-1.253
20	0.539	-1.567	0.626	-1.591	0.039	-1.255
25	0.486	-1.594	0.566	-1.617	0.035	-1.262
30	0.446	-1.613	0.520	-1.634	0.032	-1.267
35	0.415	-1.626	0.484	-1.648	0.029	-1.271
40	0.389	-1.636	0.455	-1.658	0.028	-1.274
45	0.368	-1.645	0.430	-1.665	0.026	-1.276
50	0.350	-1.651	0.409	-1.672	0.025	-1.278
55	0.334	-1.656	0.390	-1.677	0.024	-1.279
60	0.320	-1.661	0.374	-1.681	0.023	-1.280
65	0.308	-1.665	0.360	-1.685	0.022	-1.281
70	0.297	-1.668	0.348	-1.688	0.021	-1.282

σ is posterior standard deviation. μ is posterior mean, assuming prior is $\bar{\mu}_H + \mu_{0_H}$ and all experience signals are μ_H.

Table 10: Policy Experiments

price menu	average revenue per household per week	market share
actual menu with 3 plans	\$3.656	23.18
eliminate the low fee plan	2.676	15.12
eliminate the middle fee plan	4.105	23.17
eliminate the high fee plan	2.757	23.15
offer only the low fee plan	2.762	23.11
offer only the middle fee plan	1.349	11.41
offer only the high fee plan	2.665	11.29