Confessions of an Internet Monopolist: Demand Estimation for a Versioned Information Good

Chappell, Henry; Guimaraes, Paulo and Ozturk, Orgul
University of South Carolina, University of South Carolina, University of South Carolina

2006

Confessions of an Internet Monopolist:
Demand Estimation for a Versioned Information Good

Henry W. Chappell, Jr.
Department of Economics
University of South Carolina
chappell@moore.sc.edu

Paulo Guimarães
Department of Biostatistics, Bioinformatics and Epidemiology
Medical University of South Carolina
guimarap@musc.edu

Orgul Ozturk
Department of Economics
University of South Carolina
ozturk@moore.sc.edu

Abstract

We investigate profit-maximizing versioning plans for an information goods monopolist. The analysis employs data obtained from a web-based field experiment in which potential buyers were offered information goods in varied price-quality configurations. Maximum simulated likelihood (MSL) methods are used to estimate parameters describing the distribution of utility function parameters across potential buyers of the good. The resulting estimates are used to examine the impact of versioning on seller profits and market efficiency.

Keywords:
Versioning, price discrimination, field experiment, maximum simulated likelihood

JEL Categories:

C81 - Methodology for Collecting, Estimating, and Organizing Microeconomic Data
C93 - Field Experiments
D12 - Consumer Economics: Empirical Analysis
D42 - Monopoly
D83 - Search; Learning; Information and Knowledge

June 12, 2007

Confessions of an Internet Monopolist: Demand Estimation for a Versioned Information Good

The Internet has spurred the creation of new markets and changed the nature of competition in others. The ability of consumers to quickly and efficiently search over prices suggests that competitive pressures will be intense in retail markets with more than a few sellers of homogeneous goods. However, in markets that are monopolized or where product differentiation gives sellers protection from competition, the Internet provides sellers new opportunities to exploit market power through price discrimination. By tracking the online buying patterns of customers, sellers may be able to learn about buyer attributes and extract more of the surplus generated by trade.[1]

Price discrimination may be particularly important in information goods markets. Information goods are those that can be produced in a digital format; this frequently makes them suitable for sale and distribution via the Internet. Examples include textual matter, music and video entertainment, and software packages. Varian and Shapiro (1999) note that the production of information goods usually involves high fixed costs of development, but low marginal costs of reproduction. For example, the development of the first copy of the Windows XP operating system was very costly, but the cost of producing an additional copy on CD-ROM is close to zero. Information goods are not likely to be produced and sold in competitive markets. Competition drives prices toward marginal costs, but marginal cost pricing will not generate revenue adequate to cover the fixed costs of development. Instead, pricing is likely to be "value-based" (Varian and Shapiro, 1999). Successful producers of information goods will learn to extract more revenue from those who are willing to pay more.

Varian (2000) has shown that product versioning can, in effect, serve as a means of price discriminating in markets for information goods. By their nature, information goods are collections or bundles of data. Once an information good has been produced, it is possible to produce variants that differ in terms of the included content.[2] For example, the "deluxe" version of a software package can be transformed into a "standard" version by the subtraction of features. Even if sellers cannot distinguish customers by "willingness-to-pay" *a priori*, appropriate versioning and pricing schemes can induce buyers to sort themselves via their purchase decisions. The result can be higher revenue and profit for the seller.

---

[1] The problem of optimal price experimentation by a monopolist facing uncertain demand has been explored by Rothschild (1974) and Aghion, Bolton, Harris, and Julien (1991). Loginova and Taylor (2005) investigate incentives for monopoly experimentation when sellers can charge personalized prices. They find that learning by the monopolist is often thwarted by strategic behavior of buyers. Aquisti and Varian (2003) report similar results but also characterize circumstances in which conditioning prices on past purchase behavior can be profitable. Esteves (2005) finds that under duopoly, firms may eschew learning about the loyalties of individual buyers because subsequent price discrimination can lead to a more aggressively competitive outcome.

[2] The literature on information goods frequently refers to issues related to the "bundling" or "aggregation" of component goods (Bakos and Brynjolfsson, 2000). Versioning can often be though of as a special instance of bundling in which lower quality versions (bundles) can be formed by subtracting components from versions (bundles) of higher quality.

In this paper, we examine several aspects of product versioning. In section I, we review the analysis of Varian (1997) and extend his work by further characterizing demand conditions under which versioning will be profitable. Section II describes a general econometric method that can be used to estimate parameters of demand systems for versioned information goods. We have applied this method to an original data set that describes the behavior of potential buyers who were given opportunities to purchase versioned information goods (bundles of personalized digital images) in an online environment.[3] Our data set was generated in a field experiment carried out by the authors using a "real-world" Internet business; the design of those experiments is described in section III of the paper.[4] In section IV, econometric estimates are presented and discussed. Using these estimates, we design profit-maximizing versioning plans for the good, and we calculate how versioning affects seller profits and market efficiency. Conclusions follow in section V.

I. Theoretical Background

In this section we begin by presenting a simple diagrammatic illustration of the versioning problem. This follows the example of Varian (2000). We then further consider the problem of characterizing conditions under which versioning might be profitable.

Optimal Versioning: A Diagrammatic Approach

Following Varian (2000), we assume that there are two groups of customers (potential buyers) who differ in their willingness to pay for quality embodied in a single unit of an information good. We refer to these groups as "type 1" (low-demand) and "type 2" (high-demand) customers. Figure 1 illustrates the marginal valuation (of quality) schedules for individual members of these groups. Individuals' total valuations are measured by the usual areas beneath the marginal valuation schedules. We assume that production of the good is costless, regardless of quality level.

Initially suppose that the single seller of the good can identify customers by type. For this case, type 1 customers would each be charged a price P_1 equal to area A for a version of the good with quality level S_1 and type 2 customers would be charged a price P_2 equal to area A+B+C for a version with quality level S_2 (in each case, customers are charged a price equal to total willingness to pay).

Now suppose that the seller cannot observe a customer's type. The seller might again consider producing versions with quality levels S_1 and S_2 to be sold at prices $P_1 = \text{A}$ and $P_2 = \text{A} + \text{B} + \text{C}$, hoping to extract all surplus from buyers. However, type 2

---

[3] The Loginova-Taylor (cf. note 1) results suggest that one should be concerned with possible strategic behavior of buyers in such experiments. However, our experimental scenario was not one in which repeat buying or personalized pricing was either planned by the seller or likely to be expected by buyers.

[4] Harrison and List (2005) provide a survey of field experiments in economics and provide a comparative discussion of field and laboratory experiments.

customers would not choose the bundle intended for them. By choosing the low-quality bundle, each high-demand customer can earn surplus equal to area B rather than a surplus of zero. Therefore, under this plan, the seller would earn only A from each customer.

However, the seller can do better. Suppose that the seller again charges P_1=A for a low quality good, but now charges a price P_2=A+C for the high quality good. Type 2 buyers are just willing to buy the high quality version, since they now retain surplus B from the purchase of either version. The seller receives A from each type 1 buyer and A+C from each type 2 buyer.

From the seller's viewpoint, this is an improvement, but is still not optimal. Suppose that the seller slightly reduces the quality of the low-quality version, as depicted in Figure 2. This reduces the revenue that can be obtained from a type 1 customer by the amount of the shaded triangle. However, because it makes the low quality version less attractive, it increases the willingness of type 2 customers to pay for the high quality version. The amount of the increase is measured by the area of the black trapezoid and, as the diagram shows, the gain from type 2 customers can exceed the loss from type 1 customers. To select an optimal version, the seller would reduce the quality of the low-quality version to the point where the marginal reduction in revenue from type 1 customers just equals the marginal gain in revenue from type 2 customers. For the case where there are equal numbers of the two customer types, that occurs at quality level S_1^* in Figure 2.

When is Versioning Profitable?

The preceding section demonstrates that versioning can be profitable when customers differ in their valuations of quality. However, the example relies on a special case where there are two distinct customer types with widely differing preferences. In this section we explore conditions required for the existence of a profitable versioning scheme for a particular class of buyer utility functions that we later employ in our empirical analysis. The results serve to emphasize that the preceding example from Varian is a very special one.

We begin by considering the choice problem facing a potential buyer. The buyer has a utility function, $U = VS^b$, which associates a gross valuation, *U*, (measured in "dollars worth" units) with a versioned information good that has quality level S.[5] We assume that the utility function parameters satisfy the conditions $V \geq 0$ and $0 \leq b \leq 1$.[6] Versioning will be profitable only when customers are not identical. We initially assume that customers differ only in terms of the parameter *V*, which is randomly distributed across customers according to a twice differentiable cumulative distribution function,

---

[5] This particular utility function form is employed later in our empirical work; however, the theoretical results presented in this section generalize to utility functions of the form $U = Vf(S)$, where $f'(S) > 0$.

[6] These assumptions assure that the marginal utility of additional quality is non-negative and non-increasing. The latter condition ($0 \leq b \leq 1$) is not necessary to establish the results presented in this section, but it does rule out other implausible implications; cf. note 35.

$G(V)$. The number of consumers is normalized to equal one and production costs are again assumed to equal zero.

Consider first the case where the seller offers a single version of the good for sale. This version has quality *S* and is sold at a price *P* selected by the seller. For given *S* and *P*, an individual will buy if $VS^b > P$, or, equivalently, $V > P/S^b$. The fraction of customers satisfying the latter condition is given by $1-G(P/S^b)$, so the seller's profit is:

$$\Pi = P\left[1-G(P/S^b)\right]$$

The first order condition for this problem requires that the profit-maximizing price, P^*, be chosen to satisfy:

$$\frac{\partial \Pi}{\partial P} = \left[1-G\left(\frac{P^*}{S^b}\right)\right] - \frac{P^*}{S^b} g\left(\frac{P^*}{S^b}\right) = 0$$

or

$$h(T^*) = \frac{1}{T^*}$$

where $T^* = P^*/S^b$, and *h* and *g* respectively denote the hazard and density functions associated with the cumulative distribution *G*.[7] The second order conditions are satisfied provided that $2g(T^*) + T^* g'(T^*) > 0$.

We next consider the case where the monopolist offers two different versions of the good. Let S_1 be the quality of version 1 (the lower quality version), let S_2 be the quality of version 2 (the higher quality version), and let P_1 and P_2 be prices for those two versions. Each buyer maximizes utility by choosing to purchase either of two versions of the good, or neither. His/her choice may be summarized by the following set of conditions:

If $VS_1^b - P_1 \leq 0$ and $VS_2^b - P_2 \leq 0$, do not purchase; (1a)

If $VS_1^b - P_1 > 0$ and $VS_1^b - P_1 \geq VS_2^b - P_2$, purchase version 1; (1b)

If $VS_2^b - P_2 > 0$ and $VS_2^b - P_2 > VS_1^b - P_1$, purchase version 2. (1c)

[7] The hazard function is the ratio of the probability density function to the survival function, $g/(1-G)$.

These conditions imply that a customer chooses the product version that yields greatest valuation net of price, or chooses not to buy if neither version offers a positive net valuation. Rewriting the above conditions in a more compact form we have:

If $V \le T_1$ and $V \le T_2$, do not purchase;

If $V > T_1$ and $V \ge T_{12}$, purchase version 1;

If $V > T_2$ and $V > T_{12}$, purchase version 2,

where $T_1 = P_1 / S_1^b$, $T_2 = P_2 / S_2^b$, and $T_{12} = \left(P_2 - P_1\right) / \left(S_2^b - S_1^b\right)$.

If versioning is to be a profitable strategy for the seller, then prices and versions must be selected so that there are positive sales of both versions. This requires that $T_{12} > T_2 > T_1$ or, in terms of prices and quality, that $P_2 S_1^b > P_1 S_2^b$. The rationale for this requirement can be illustrated with a graphical argument. In Figure 3 we show the (positive) net utility gained by a buyer as a function of the parameter *V*. The steepest line (with slope S_2^b) represents the net valuation obtained when the buyer consumes the higher quality version of the good, while the more gently sloped line (with slope S_1^b) represents the net valuation obtained from the lower quality good. The point where the lines intersect identifies a buyer whose value of *V* leaves her indifferent to versions 1 and 2. In the diagram, we illustrate the case where some buyers select each version, which clearly requires that $T_{12} > T_2 > T_1$. Given this condition, the profit function for a versioning monopolist is:

$$\Pi = P_2[1 - G(T_{12})] + P_1[G(T_{12}) - G(T_1)]$$

and first order conditions are:

$$\frac{\partial \Pi}{\partial P_1} = \frac{P_2}{S_2^b - S_1^b} g\left(T_{12}\right) + \left[G\left(T_{12}\right) - G\left(T_1\right)\right] - \frac{P_1}{S_2^b - S_1^b} g\left(T_{12}\right) - \frac{P_1}{S_1^b} g\left(T_1\right) = 0$$

$$\frac{\partial \Pi}{\partial P_2} = \left[1 - G\left(T_{12}\right)\right] - \frac{P_2}{S_2^b - S_1^b} g\left(T_{12}\right) + \frac{P_1}{S_2^b - S_1^b} g\left(T_{12}\right) = 0,$$

or, after simplifying:

$$\frac{\partial \Pi}{\partial P_1} = T_{12} g\left(T_{12}\right) + G\left(T_{12}\right) - G\left(T_1\right) - T_1 g\left(T_1\right) = 0 \qquad (2a)$$

$$\frac{\partial \Pi}{\partial P_2} = 1 - G(T_{12}) - T_{12} g(T_{12}) = 0. \tag{2b}$$

Assuming an optimum exists, and letting asterisks denote values at the optimum, the latter condition simplifies to:

$$h(T_{12}^*) = \frac{1}{T_{12}^*}. \tag{3b}$$

Using condition (2b) we know that $1 = G(T_{12}^*) + T_{12}^* g(T_{12}^*)$. Substituting this expression into (2a) and rearranging we obtain:

$$h(T_1^*) = \frac{1}{T_1^*}. \tag{3a}$$

Note that conditions (3a) and (3b) require that the hazard function associated with the distribution G intersect the curve $1/x$ at two distinct points. Also, recall that the solution we obtained for the problem of a non-versioning seller also occurred at a point of intersection between h and $1/x$. That point, T^*, is distinct from T_1 and T_{12}. Therefore, for the case under consideration, a profitable versioning plan can exist only if the hazard function associated with G intersects the curve $1/x$ in at least three different locations.

Well-known probability distributions fail to satisfy this requirement. These include all distributions with non-decreasing hazard rates[8] (e.g., uniform, normal, logistic, extreme-value, chi-squared, Laplace, exponential distributions, and the Gamma and Beta distributions for a range of values of the parameters) as well as distributions with non-increasing hazard rates (e.g., the Pareto distribution and the Gamma distribution for a range of values of the parameters). It is possible for hazard functions for other distributions, including the lognormal and Beta distributions, to intersect the function 1/x in three distinct points, but only for specific parameter values. We conclude that for utility functions of the form $U = VS^b$, if V is random but b is not, then versioning will be profitable only when very special restrictive assumptions are made about the distribution of V.[9]

Anticipating our subsequent empirical work, we can state this result a bit differently. For utility functions of the form $U = VS^b$, if V is uniformly distributed and b is non-random, then no profitable versioning plan exists. However, we can show by example that if b is random, then profitable versioning *is* possible. Consequently, a necessary condition for profitable versioning is that b be random. This has an implication

---

[8] All log-concave distribution functions have increasing hazard functions.

[9] Jing (2000) reports similar results, finding that versioning is not profitable when g is log-concave and $2g(x) + xg'(x) > 0$. Varian's versioning example (replicated in section I of this paper) assumes that G is not a continuous function.

for empirical work—if we wish to investigate profitable versioning schemes, model specifications should be sufficiently general to permit randomness in both *b* and *V*.

II. Econometric Method

Consider a sample of customers indexed by $i = 1,..., N$, each with a utility function of the form $U_i = V_i S^{b_i}$. Customers differ in their utility function parameters V_i and b_i , which are randomly distributed over the population. Let $G(V_i, b_i, \mathbf{?})$ be the joint distribution function from which each individual's utility function parameters are drawn, where $\mathbf{?}$ is a vector of parameters characterizing that distribution. Further, assume that the purchase options that customers face are individual-specific. Each customer is presented with two versions of a good, with prices and quality levels given by P_{1i}, P_{2i}, S_{1i} and S_{2i}. Utility maximizing purchase choices are determined as described in conditions (1).

Let R_{0i} be the probability that an arbitrary customer *i* chooses not to buy, given his choice opportunities (described by P_{1i}, P_{2i}, S_{1i} and S_{2i}) and given the distribution from which his utility function is drawn. Similarly, let R_{1i} and R_{2i} be the probabilities that customer *i* purchases versions 1 and 2, respectively. These probabilities depend upon the functional form of the distribution *G*, the parameters of the distribution, and the prices and qualities of the goods available for purchase by customer *i*. In equation form:

$$R_{0i} = R_0\left(P_{1i}, P_{2i}, S_{1i}, S_{2i}, \mathbf{?}\right) \tag{4a}$$

$$R_{1i} = R_1\left(P_{1i}, P_{2i}, S_{1i}, S_{2i}, \mathbf{?}\right) \tag{4b}$$

$$R_{2i} = R_2\left(P_{1i}, P_{2i}, S_{1i}, S_{2i}, \mathbf{?}\right). \tag{4c}$$

As econometricians, we observe the actual choices made by buyers in the sample, as well as the options they chose from. Our task is to estimate the parameters $\mathbf{?}$; to do so we wish to employ the maximum likelihood method. Denoting the observed purchase choices for buyer *i* by the dummy variables C_{0i}, C_{1i}, and C_{2i}, respectively indicating no purchase ($C_{0i} = 1$), purchase of version 1 ($C_{1i} = 1$), and purchase of version 2 ($C_{2i} = 1$), the likelihood function for an individual observation is given by:

$$L_i = R_{0i}^{C_{0i}} R_{1i}^{C_{1i}} R_{2i}^{C_{2i}}. \tag{5}$$

The likelihood for the sample of observations is the product of the individual likelihoods:

$$L = \prod_{i=1}^{N} L_i. \tag{6}$$

The maximum likelihood method requires that we find values for **?** that maximize *L*. A number of algorithms are available for this numerical task.

A difficulty with the estimation procedure outlined above is that it is not possible to derive closed-form solutions for the probability functions R_0, R_1, and R_2 in conditions (4) for all possible specifications for *G*. Moreover, it is not empirically desirable to limit forms for *G* to those for which analytical solutions for the probabilities are possible. However, it is possible to use the maximum likelihood method employing numerical rather than analytical calculations of probabilities. In the remainder of this section, we describe our implementation of a "maximum simulated likelihood" (MSL) method for estimation of our model.[10]

Consider customer *i*, who accounts for a single observation in a sample of size *N*. Also, consider arbitrarily selected values for **?**, denoted $\mathbf{?}_0$. From the distribution $G(V_i, b_i, \mathbf{?}_0)$, draw a random sample of simulated observations, $j = 1,...,N_s$, for V_j and b_j. For each V_j and b_j drawn for customer *i*, determine what his choice would be when facing prices and versions P_{1i}, P_{2i}, S_{1i} and S_{2i}. This is done by applying conditions (1) for each observation j. This step yields a series of simulated choice outcomes, C_{0j}, C_{1j}, and C_{2j}, for $j = 1,...,N_s$. Given these series, we can approximate the probability that customer *i* will make choices in each of the three categories with the frequencies observed in the simulated sample. That is

$$R_{0i} \approx \frac{\sum_{j=1}^{N_s} C_{0j}}{N_s},\ R_{1i} \approx \frac{\sum_{j=1}^{N_s} C_{1j}}{N_s},\ \text{and}\ R_{2i} \approx \frac{\sum_{j=1}^{N_s} C_{2j}}{N_s}.$$

For sufficiently large values of N_s, these approximations approach any desired level of accuracy. Using these values for R_{0i}, R_{1i}, and R_{2i}, one can calculate the likelihood for observation *i* given parameters $\mathbf{?}_0$ using equation (5).

Now repeat the procedure described in the preceding paragraph for each observation in the original sample, that is for $i = 1,...,N$. Doing so yields a value for the likelihood for each observation, and, using equation (6), a value of the likelihood function for the sample. We are therefore able to evaluate the likelihood function for arbitrary

[10] Discussions of maximum simulated likelihood estimation are provided by Arias and Cox (1999), Gourieroux and Monfort (1993), Hajivassiliou and Rudd (1994), Lee (1995), Lerman and Manski (1981), Stern (1997), and Train (2003). As Train (2003, 242) notes, the approximation of probabilities in the MSL procedure introduces a bias in estimation; however, the bias diminishes as the number of draws used in simulation increases. Arias and Cox (1999) further note that when probabilities are approximated by frequencies, the simulated probabilities will not be continuous functions of the underlying parameters, and standard optimization algorithms for maximum likelihood estimation may fail. The use of large numbers of draws in simulation increases the smoothness of the simulated likelihood function and improves the performance of numerical optimization methods.

parameter values. The only remaining problem is to find parameter values that maximize the likelihood function, but conventional methods can be employed for this purpose.[11]

III. A Market Experiment

This section describes a market experiment we have undertaken to investigate the demand for a versioned information good. The experiment was specifically designed to generate data suitable for the estimation method described in Section II.

Introduction

The econometric method that we have described is very general; because it is simulation-based, there are essentially no restrictions imposed on permitted utility function forms or on the distributions from which utility function parameters are drawn. Data requirements are somewhat more demanding. We have assumed that we observe a sample of individual consumers' choices, and that these consumers have been presented with varied price and quality options. For information goods sold via the Internet, such data may not be difficult for sellers to gather. For example, Amazon identifies its regular customers when they visit the site, keeps track of their purchase histories, and offers personalized suggestions for shopping. It would not be difficult to offer customers personalized prices (perhaps in the form of discount offers) on particular products in order to learn about demand characteristics.[12]

For academic researchers, obtaining such information is more difficult. Firms may be willing to experiment with prices and versions, but they are not likely to publicize their experiments or release the data they produce. The data employed in this paper were generated in a field experiment carried out by an existing Internet business, Homework Hero (www.homeworkhero.com). Homework Hero is operated by one of the authors of this paper (Chappell), who undertook the market experiment for the express purpose of obtaining data for academic use.[13]

Homework Hero is a service offered to K-12 schools that permits teachers to post homework assignments for their students on the web.[14] Each teacher using the service maintains a web page on the Homework Hero site and can update assignments posted there through a web-based form. The service provided by Homework Hero is normally purchased by schools rather than by individual teachers. However, the market experiment

---

[11] As we describe later (cf. footnote 33) we have modified the method described here slightly to take advantage of special features of our data and our empirical model.
[12] In fact, Amazon has carried out pricing experiments. In the summer of 2000 DVD prices were varied randomly to visitors in an experiment intended to provide information about customer demand (Wolverton, 2000).
[13] Normally, when a firm undertakes a pricing experiment, it must weight the value of information gained against profits lost while offering suboptimal experimental prices. An advantage to being author-owners is that we were not concerned with revenues generated by the experiment. We agreed in advance to donate all revenues generated by the experiment to the Economics Department at the University of South Carolina.
[14] In September 2004, about 200 schools subscribed to the service, and about 5000 teacher pages existed. The site received about 100,000 unique visits and about 1.2 million "total hits" per week.

involved selling a complementary product, a collection of personalized digital images, directly to teachers.

Experimental Design

The Homework Hero web site is organized so that each subscribing school has a set of school-specific pages and a database that holds the content of assignment pages for the school's teachers. Within a school, each teacher's assignment page uses a common template that determines the overall format of the page, but individual teachers provide page content. Teachers have options that permit them to post images and use HTML (Hypertext Markup Language) to create a personalized appearance.[15] In practice, many teachers go to some effort to customize the appearance of their pages with both HTML formatting and the display of images.

In our experiment, we sold bundles of images for teachers to display on their assignment pages.[16] Each of the images in a bundle included a graphic rendering of the teacher name, with the style, color scheme, animation, and theme of the images varying. Essentially, the image collections offered opportunities for personalization of a teacher page. Having multiple images would permit a teacher to vary the appearance of the page over time. Some of the bundles offered included images with seasonal themes, so it would be natural to display them in sequence over the academic year. Hereafter, we will assume that the quality of a bundle of images is given by the number of images it contains.[17] The appendix to the paper displays several images from the collections offered for sale.

Homework Hero has some market power in the sale of these image collections. The images sold were customized for use on Homework Hero and were personalized for buyers. Further, for the images to be displayed on a page, they must first be stored on the web, and Homework Hero provided storage space for the images sold. While it is possible for users to display other images on assignment pages, it would have been difficult for a teacher to replicate the product sold in the experiment.

A total of 38 subscribing schools were selected for market experiments.[18] The selected schools generally had large numbers of teachers actively maintaining assignment

[15] There are actually two vintages of the Homework Hero server software that offer different options for displaying graphics. In the early vintage, teachers can only display graphics that have been previously posted elsewhere on the web. In the more recent vintage, teachers can upload a single graphic from their home computer to Homework Hero for display. Both software vintages were in use by schools included in the market experiment.

[16] Images were created by the authors using two software packages: Paint Shop Pro and Ulead Cool 3D.

[17] Because we know precisely which images were offered in each bundle, in principle we could investigate whether specific images were especially highly valued. Because overall sales rates were low, we have not attempted to refine the empirical models in this fashion.

[18] Some schools subscribe to a "non-commercial" version of the service that normally excludes all advertising from assignment pages. Because our experiments displayed ads viewed by teachers, site managers at these schools were notified in advance that ads would be displayed during a brief experiment undertaken for academic purposes. However, teachers viewing the ads would typically not have known that the sales offer was associated with a market experiment.

pages, but in other respects were similar to subscribing schools that were not selected. The specific image bundles and prices offered to teachers were common across teachers at a school, but varied across schools. There were two reasons for this. First, given the design of the Homework Hero web site, it was technically easy to customize offers to schools, but would have been difficult to vary offers to individual teachers.[19] Second, we considered it undesirable to have teachers know that prices and options were being varied, and this would have been more apparent if offers were varied across teachers within a school.[20]

At each school, the sale of images took place over an eight-day period, starting at 9:00 pm on a Sunday and ending at 9:00 pm on a Monday.[21] All school experiments took place in the fall of 2004, however the experiments were not all conducted simultaneously. Experiments were run at several schools in every week in the interval from August 21 to October 25.[22] At each school, teachers could choose to purchase either of two image bundles (or neither). The larger bundle always included all of the images in the smaller bundle, plus additional ones. The larger bundle was also always offered at a higher price than the smaller, however, per-image prices varied across versions and schools. Bundle sizes ranged from 1 to 12 images, and prices ranged from $0 to $10 across experiments; Table 1 provides additional summary data on the bundle sizes and prices offered, the number of teachers receiving offers, and the numbers buying high and low quality versions.[23]

Apart from the specific offers, the mechanics of the experiments were the same across schools. Normally, when a teacher updates assignments on Homework Hero, he or she sees a screen that confirms a successful update and offers several options (view the updated page, update again, or follow links to other pages). When experiments were in progress, this screen also displayed an ad describing the image sales offer. The ad showed an example of a personalized graphic offered for sale and briefly described the sales options available (including numbers of images and prices for the offered collections). The ad also included hyperlinks to a more elaborate sales information page where teachers could browse through the images offered for sale and learn details of the product

[19] In addition, Homework Hero's privacy policy rules out the use of cookies that would identify an individual for purposes of offering an individual-specific price.
[20] Knowledge that prices varied could have caused some confusion and/or resentment among customers. Since Homework Hero is a "real" business, we wanted to avoid this possibility. In addition, the experiment should attempt to create conditions similar to those that would prevail in a post-experimental selling stage in a real business. Once an optimal versioning and pricing scheme is established by a firm, all customers would be presented with identical options. Because we offered uniform prices and options within schools, our experimental customers were placed in a similar environment.
[21] Experiments involving two schools were intentionally extended by a day because the normal end date fell on the Labor Day holiday. One experiment was inadvertently extended by 2 days because of an oversight by the authors.
[22] The work associated with selling and producing the products would have made it difficult for the authors to manage 36 experiments simultaneously. The staggering of school opening dates (from early August to mid-September) also led us to vary start dates for the experiments. We conducted the experiments early in the school year because teacher usage is heavy at this time.
[23] The number of distinct offers (35) is less than the number of schools (38) because several small schools were combined in identical treatments.

offer. Both the initial ad and the more detailed offer page made it clear that the offer would expire on a specified date. The detailed offer page indicated that group orders could not be accepted; only individual teachers were permitted to buy. The ad was altered to display a different sample image twice during each experiment. On the last day of the experiment, a bold red message emphasized that the sale would end at 9:00 pm that night. Throughout the experiment the sales offer was presented as ordinary commercial activity, without any mention of an experimental context.[24] Teachers who purchased images retained the rights to employ those images on Homework Hero as long as their schools continued to subscribe to the service.

When a teacher decided to buy, the order was placed by submitting a form on the web. Teachers could either pay immediately via credit card (using the PayPal transaction service) or pay later after we sent a bill in the mail.[25] All teachers who ordered images eventually paid for them. Teachers who chose to purchase the smaller of the two available bundles were told that they could upgrade to the larger version at anytime within the experimental period; however, only one buyer chose to do so.

Once an order was completed, the authors produced the image bundles by editing the displayed name in previously designed image templates. The images were stored on the Homework Hero server and purchasers were notified of the URLs (web addresses) for the images. By pasting a URL into the appropriate entry box on the usual Homework Hero assignment update form, a teacher could display a image on his assignment page. Later, by changing a single number in the URL pasted in the form, a teacher could display an alternative image from the collection.

Summary Data

Across all participating schools, a total of 1440 teacher assignment pages were updated in the course of the experiment.[26] We could identify the owner of each of those pages, and therefore could determine the identities of all teachers who viewed the initial ad describing the sales offer.[27] Of these 1440 teachers, ultimately 33 teachers purchased image collections, including 11 who “purchased” bundles that were priced at zero. Total revenue from the experiment’s sales was $62.00.

Obviously, purchases were infrequent, but the low sales frequency should not be surprising. Teachers do not come to Homework Hero intending to make purchases; i.e., they are not shoppers, *a priori*. The offer they see is in the form of a banner ad,

---

[24] An exception has been noted (cf. note 18), however, most teachers would not have known about the experimental nature of the offer.

[25] Because teachers often update pages in short periods of time between class periods, we felt that it would be helpful to keep the purchase process as simple as possible. Since credit card forms are sometimes cumbersome, we offered the “pay later” option.

[26] Pages for clubs, sports teams, other organizations, or pages maintained by groups of more than two teachers were excluded from the sample. None of the individuals maintaining these pages ordered image bundles during the experiment.

[27] Unfortunately, we were not able to unambiguously determine which teachers clicked to proceed to the detailed sales offer page.

augmented with a brief textual message. Click-through rates on banner ads displayed on the web are generally low—according to a March 2000 report, the Internet-wide average click-through rate was 0.38%, and sales rates are necessarily lower than click-through rates.[28] With a sales rate at about 2% of the overall audience, our results substantially outperform the Internet average for banner advertising.[29] This also is not surprising—the product was targeted to an audience that should have been receptive (since all ad recipients were already Homework Hero users). Nevertheless the paucity of sales limits the scope of the empirical investigation that can be undertaken.

Using the data obtained from experiments, our purpose is to devise a profit-maximizing versioning and pricing plan. It is a characteristic of versioning schemes that prices *not* vary across identifiable groups or individuals. Instead, all individuals are presented with identical options and they sort themselves through their purchase decisions. Given our intentions, we purposely refrain from using individual- and group-specific data in estimation of the demand system.

For other purposes, having knowledge of how individual teacher characteristics affect purchase decisions could be useful, and, in the course of our work, we have collected data to describe some of these characteristics.[30] For example, teacher assignment pages usually reveal grade levels and subject areas taught, and whether a teacher was male or female. Assignment pages also give some quantifiable indications of the willingness of teachers to produce a customized web page. For example, we can observe whether teachers had posted images or used customized HTML on their page prior to the experiment. We can also obtain data describing attributes of the school where each teacher was employed (student population, grade levels, public or private status, religious affiliations, etc.) and demographic information for geographic areas in which specific schools are located.

In the following section, we will present econometric results that employ the methods described in section II. However, as a preliminary empirical exercise, we have estimated a simple conditional logit model using our data. This model specifies that customer choices depend upon the qualities and prices of the offered bundles. The results in Table 2 reveal that both price and quality (number of images) are related to customer choices in an expected fashion—a lower price and a higher quality level make a bundle more attractive and therefore increase the probability that it will be purchased. Coefficients for both price and quality differ significantly from zero in the appropriate directions. While the logit specification is inappropriate for our purposes, it is reassuring

---

[28] See http://www.consult-x.com/papers/webpromotion-bannerads.htm. Hoffman and Novak (2000) quote an earlier (1999) CyberAtlas report showing an average click-through rate of 0.58%.

[29] More than 10% of the target population clicked-through the initial ad at some point during the experiment (based on 156 recorded "hits" on the sales page banner and a population of 1440 teachers who updated pages in the experimental period). About one-fifth of those who clicked through ultimately purchased.

[30] For example, one might wish to design third degree price discrimination schemes as well as versioning schemes.

to see that price and quality exhibit expected impacts on customer choices when this well-known empirical model is estimated.[31]

IV. Econometric Results

Using the method described in Section II, we will estimate parameters characterizing the distribution of individual teachers' utility functions. Each teacher *i* is assumed to have a utility function of the form $U_i = V_i S^{b_i}$, where V_i and b_i are random variables drawn from a distribution $G(V_i, b_i, \mathbf{?})$. The quality of a bundle, *S*, is measured by the number of images it contains.[32] Although it would be possible to employ any of a wide range of distributions for *G*, we begin by using a specific simple distribution. We assume that V_i is uniformly distributed from 0 to $V_{\max}$ ($V_{\max} > 0$) but with a mass of probability at $V_i = 0$. The probability that $V_i = 0$ is indicated by parameter M_0. We also assume that b_i is uniformly distributed from b_{lo} to b_{hi} (where $0 \le b_{lo} < b_{hi} \le 1$) and that the distributions for V_i and b_i are independent. Our problem is to find MSL estimates of the parameters $V_{\max}$, M_0, b_{lo}, and b_{hi}.[33] To insure accuracy of our estimates and smoothness of the simulated likelihood function, we set $N_s = 1{,}000{,}000$ (in evaluating the likelihood function, N_s is the number of simulated observations used in calculating a probability for each observation in the original data).[34]

Two sets of model estimates are presented in Table 3. The estimations differ only in terms of restrictions imposed on the coefficients. Model 1 imposes no restrictions on coefficient values, but yields estimates of b_{lo} and b_{hi} that are outside of theoretically specified limits (the estimated value for b_{lo} is less than zero and that for b_{hi} is greater

[31] Most importantly, the logit model apparently precludes profitable versioning (we have found no profitable versioning plans and conjecture that this is a general result). In addition, the random utility specification underlying the logit model contains only a single an additive random error term that varies across choices as well as individuals. This assumption permits some peculiar outcomes in the context of our application. For example, a buyer might choose a smaller bundle when offered a larger bundle (that includes the smaller as a subset) even when the larger bundle is offered at a lower price.

[32] Because we have recorded more detailed information on the composition of bundles, it might be possible to be more general in defining quality. However, given the limited number of observed purchases, we have not pursued this possibility.

[33] Our estimation procedure for this application differs slightly from that described earlier in order to take advantage of special features of our model and data. First, when there is a mass point of probability for $V_i = 0$, it is only necessary to simulate choices over the range of positive values for V_i. Second, because all price and version options were identical across teachers in a school, and because we do not include any individual-specific variables, there are a small number of possible values for the likelihood function for any observation. It is not necessarily to simulate for each observation separately, since many observations within a given school are identical from an econometric perspective.

[34] We employed the MLPROC procedure in TSP 4.5 to implement this method.

than one).[35] Although the point estimates are outside the bounds suggested by theory, neither differs significantly from its theoretically prescribed limit.[36] Model 2 imposes the theoretical restriction that $0 \le b_{lo} < b_{hi} \le 1$ (by setting $b_{lo} = 0$ and $b_{hi} = 1.0$). In principle, it would be possible to further restrict the model so that $b_{lo} = b_{hi}$, implying that b_i is non-random (and, from Section I, that no versioning scheme is possible). However, imposition of this restriction leads our model to imply that some observed choice patterns should occur with probability zero, and we therefore rule out this case.[37]

Although the uniform distributions assumed for V_i and b_i are plausible, it would be desirable to adopt a more general specification. The Beta distribution is an attractive option for this purpose for two reasons. First, it is flexible. Depending on parameter values, the beta probability density function can have any of a variety of shapes. Second, the uniform is a special case of the Beta distribution, so the hypothesis that a distribution is uniform can be tested as a restriction in a more general model.

Uniform and Beta distributions share parameters that determine upper and lower bounds, however, the Beta distribution has two additional "shape" parameters, denoted a and b. Permitting both V_i and b_i to have Beta rather than uniform distributions therefore requires that four additional parameters be estimated. When we attempted to estimate this more general specification, TSP's maximization algorithm failed to converge. We then searched over a grid of Beta distribution shape parameters, letting each of the shape parameters vary between 0.25 and 2.0 at intervals of 0.25. Table 4 reports estimates of the model for the best values of the shape parameters. Based on these results, a likelihood ratio test fails to reject the restrictions imposed by the uniform distribution for either V_i or b_i; consequently, we maintain the assumption of uniform distributions in the remainder of our work.[38]

We now focus on the parameter estimates for Model 2 in Table 3, where uniform distributions are assumed and the restrictions $b_{lo} = 0$ and $b_{hi} = 1.0$ are imposed. In this estimation, the mass point coefficient, M_0, has a value of 0.96, implying that 96% of the teachers had zero valuations for image bundles. Given that only about 2% of the sample actually purchased a bundle, this seems to be a reasonable estimate (teachers who ignored the ads, regardless of the reason, are appropriately considered to have zero valuations).[39]

---

[35] A negative value for b implies that the marginal utility of quality is negative, while a value of b greater than one implies that the marginal utility of quality is rising. In the latter case, if images could be added to a bundle at a constant cost, then the profits from selling to a single customer would be unbounded.

[36] A likelihood ratio test also fails to reject the joint hypothesis that $b_{lo} = 0$ and $b_{hi} = 1.0$; specifically, $c^2 = 1.067 < 5.99 = c^2_{.05}(2)$.

[37] For some price-bundle configurations, the model implies that one bundle should dominate the other for all customers. However, in the data, some customers choose each bundle.

[38] Treating the best shape parameters obtained in the grid search as those maximizing log*L*, the test results are: $c^2 = 3.36 < 9.49 = c^2_{.05}(4)$; we fail to reject the hypothesis of uniform distributions.

[39] This result does reveal the importance attached to getting the attention of potential customers, apart from selecting profit-maximizing versions and prices.

A buyer with median values of V_i and b_i (for buyers with V_i positive) would have a total valuation of $3.08 for a 12-image bundle. Given that all bundle sales took place at prices between $0 and $5, this result also appear to be reasonable.

Next we consider the problem of determining a profit maximizing versioning and pricing plan, given the demand system estimates for Model 2. For simplicity, we will continue to consider plans that offer just two versions of the good. In the optimal plan the large bundle necessarily includes all 12 images (given that we assume that marginal production cost is zero). We wish to determine the optimal size of the smaller bundle, S_1, and prices for both small and large bundles, P_1 and P_2.

To calculate the profitability of any plan, we again employ simulation methods. For given S_1, P_1, and P_2 we draw a large number (10,000) of customers from the estimated distribution of utility functions. For each simulated customer, we determine a purchase choice (no purchase, purchase the small bundle, purchase the large bundle) and the profit generated by that purchase choice. We then average profits across all 10,000 simulated customers and repeat this procedure for different values of S_1, P_1, and P_2. After examining all feasible plans, we determine which one yields the highest profit per customer.[40] In this exercise, we simulate only customers with positive V_i values (those with zero values for V_i would not purchase at any positive price and are irrelevant for the profit maximization calculation).

Results of this exercise, summarized in Table 5, show that the optimal versioning plan has $S_1 = 2$, $S_2 = 12$, $P_1 = \$1.28$ and $P_2 = \$4.84$, and yields profits per customer (for customers with non-zero valuations) of $1.56. To compare these outcomes to those that would prevail in the absence of versioning, we repeated the simulation described above, but assumed that only the large bundle was sold at a single "monopoly" price. Results in Table 5 show that the profit-maximizing single-version price of $4.30 produces a profit per customer of $1.42. Consequently, optimal versioning increases profit per customer by 11.3% over single-version monopoly pricing. In our experiment, actual revenues obtained from 1440 teachers who updated pages totaled $62.00. Our results imply that if we had instead set an optimal price for a single version, our expected revenue would have been $81.79. Had we imposed the optimal two-version package and pricing scheme, revenue would have risen further to $91.01.[41] These are small amounts, but similar percentage changes would imply huge benefits for large firms

Our results also have implications regarding the welfare consequences of versioning. Table 5 reports that total surplus per customer increases from $2.76 under single-product monopoly to $2.84 under versioning (an increase of 2.9%). Versioning increases profits and decreases consumer surplus compared to single-version monopoly,

[40] We consider only integer values for S_1.

[41] Of 1440 teachers, only 57.6 (4.00%) are estimated to have had positive valuations. Multiplying 57.6 by profits per customer yields these amounts.

but the profit increase is slightly greater than the loss of consumer surplus. The increase in overall welfare is associated with a large increase in the number of customers served. Table 5 shows that 51.0% of customers with non-zero valuations make a purchase under profit-maximizing versioning, while only 33.0% buy under the single-version monopoly regime. Of course many who buy under the versioning scenario receive a product of less than optimal quality.

Results from versioning and non-versioning monopoly scenarios can also be compared to a welfare maximizing outcome. Assuming marginal cost of zero, the welfare maximum would give each customer the maximum sized bundle of 12 images, resulting in a total surplus of $3.99 per customer, a welfare gain of 39.0% over optimal versioning. With or without versioning, the welfare cost associated with monopoly pricing is substantial.

V. Conclusions

Information goods are commonly sold in quality differentiated versions. With appropriately designed versions and prices, this practice can generate profits as customers sort themselves by willingness-to-pay when choosing which versions to purchase. However, to exploit profitable versioning opportunities, producers must have knowledge of the distribution of preferences across the population of potential buyers.

We have employed an econometric method that permits us to use data from pricing experiments to characterize distributions of customer preferences. We have applied the method to data obtained from an experiment in which we varied versions and prices of an information good sold via the Internet (the good consisted of a bundle of personalized digital images intended for display on web pages). For this application we were able to construct profit-maximizing versioning and pricing schemes. Profit-maximizing versioning increased expected profits to the seller by 11.3 % relative to a comparable single-version monopoly plan. Total welfare increased slightly under versioning, as more customers were served.

Versioning is a common practice, and firms who engage in versioning often have opportunities to experiment with prices in order to learn about demand. Information services and software products provide many examples of versioning, but the practice is common in other product markets as well. For example, wireless phone and cable television services come in alternative versions (i.e., service plans). Fees for plans can be varied over locations or over time and detailed information on usage patterns can be obtained from customers. The problem of how to use this information to devise service options and set prices is one that might be analyzed using methods similar to those described in this paper.

References

Acquisti, Alessandro, and Hal R. Varian, "Conditioning Prices on Purchase History," *Marketing Science*, forthcoming, 2005.

Aghion, Philippe, Patrick Bolton, Christopher Harris, and Bruno Julien, "Optimal Learning by Experimentation," *The Review of Economic Studies*, 1991, 58, pp. 621-654.

Arias, Carlos, and Thomas L. Cox, "Maximum Simulated Likelihood: A Brief Introduction for Practitioners." University of Wisconsin-Madison, Department of Agricultural and Applied Economics, Staff Paper 421, 1999.

Bakos, Yannis, and Erik Brynjolfsson, "Aggregation and Disaggregation of Information Goods: Implications for Bundling, Site Licensing, and Micropayment Systems," in Hal R. Varian and Brian Kahin, eds., *Internet Publishing and Beyond: The Economics of Digital Information and Intellectual Property* (Cambridge: MIT Press), 2000, pp. 114-137.

Esteves, Rosa Branca, "Pricing with Customer Recognition," Working paper, Oxford University, 2005.

Gourieroux, C. and A. Monfort. "Simulation-based inference: A Survey with Special Reference to Panel Data Models," *Journal of Econometrics*, 1993, 59, pp. 5-33.

Hajivassiliou, V. and P. Ruud. "Classical Estimation Methods for LDV Models Using Simulation." In R. Engle and D. McFadden, eds., *Handbook of Econometrics* (Amsterdam: North-Holland), 1994, pp. 2383-2441.

Harrison, Glenn, and John A. List, "Field Experiments" *Journal of Economic Literature*, 2004, 42, pp. 1009-1055.

Hoffman, Donna L. and Thomas P. Novak, "Advertising Pricing Models for the World Wide Web," in Hal R. Varian and Brian Kahin, eds., *Internet Publishing and Beyond: The Economics of Digital Information and Intellectual Property* (Cambridge: MIT Press), 2000, pp. 45-61.

Jing, Bing, "Versioning Information Goods with Network Externalities," in Soon Ang, Helmut Krcmar, Wanda J. Orlikowski, Peter Weill, Janice I. DeGross (Eds.): *Proceedings of the Twenty-First International Conference on Information Systems*, December 10-13, 2000 Brisbane, Australia. (Atlanta: Association for Information Systems) 2000.

Lee, L. "Asymptotic Bias in Simulated Maximum Likelihood Estimation of Discrete Choice Models." *Econometric Theory*, 1995, 11, pp. 437-483.

Lerman, S. and C. Manski. “On the Use of Simulated Frequencies to Approximate Choice Probabilities,” in C. Manski, and D. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications* (Cambridge: MIT Press), 1981, pp. 305-319.

Loginova, Oksana, and Curtis R. Taylor, “Price Experimentation with Strategic Buyers,” Working paper, Duke University, 2005.

Rothschild, M. “A Two-Armed Bandit Theory of Market Pricing,” *Journal of Economic Theory*, 1974, 9, pp. 185-202.

Stern, Steven. “Simulation-Based Estimation,” *Journal of Economic Literature*, 1997, 35, pp. 2006-2039.

Varian, Hal R. “Versioning Information Goods,” in Hal R. Varian and Brian Kahin, eds., *Internet Publishing and Beyond: The Economics of Digital Information and Intellectual Property* (Cambridge: MIT Press), 2000, pp. 190-202.

Varian, Hal R. and Carl Shapiro, *Information Rules* (Boston: Harvard Business School Press), 1999.

Wolverton, Troy. “Now showing: random DVD prices on Amazon, ” CNET news.com, September 5, 2000 (http://news.com.com/2100-1017-245326.html?legacy=cnet).

Table 1. Experimental Bundle and Price Configurations

S_1	S_2	P_1	P_2	Number of Teachers	Number Buying #1	Number Buying #2
1	5	1.00	5.00	21	0	0
1	7	1.00	5.00	22	0	0
1	8	0.00	4.00	23	2	1
1	9	0.00	7.00	19	1	0
2	6	0.00	1.00	58	0	1
2	7	0.00	5.00	26	0	0
2	9	1.00	3.00	153	0	2
2	11	1.00	3.00	77	0	0
2	12	1.00	4.00	39	1	0
2	12	3.00	10.00	42	0	0
3	6	3.00	6.00	26	0	0
3	7	2.00	4.00	30	0	0
3	8	1.00	2.00	25	0	0
3	10	0.00	1.00	18	2	0
3	10	0.00	3.00	20	0	0
3	11	2.00	4.00	54	0	3
3	12	0.00	3.00	23	3	1
3	12	2.00	8.00	13	0	0
4	8	1.00	4.00	46	2	0
4	9	2.00	5.00	86	1	0
4	10	0.00	2.00	44	1	0
4	10	0.00	1.00	20	1	2
4	11	1.00	2.00	98	1	0
4	11	2.00	4.00	34	0	0
5	8	4.00	5.00	16	0	0
5	8	5.00	8.00	35	0	0
5	10	5.00	9.00	22	0	0
5	12	2.00	4.00	20	0	1
5	12	3.00	5.00	165	1	2
5	12	3.00	5.00	13	0	0
5	12	3.00	5.00	16	0	0
6	9	3.00	4.00	32	0	0
6	11	3.00	6.00	35	0	0
7	10	1.00	2.00	31	0	0
7	12	0.00	3.00	38	1	3

Table 2. Conditional Logit Estimates
Standard errors in parentheses

Variable	Coefficient
Constant	-4.3599 (0.3012)
S	0.1753 (0.0641)
P	-0.5703 (0.1799)
log*L*	-172.21

Table 3. Model Estimates

Standard errors in parentheses

Parameter	Model 1	Model 2
M_0	0.9544	0.9600
	(0.0088)	(0.0068)
$V_{\max}$	1.7486	1.7968
	(0.4690)	(0.3221)
b_{lo}	-0.8925	0.0000[†]
	(0.9617)	
b_{hi}	1.3434	1.0000[†]
	(0.2436)	
$\log L$	-171.8391	-172.3724

[†]Coefficient constrained to equal the indicated value.

Table 4. Model Estimates Assuming Beta Distributions

Standard errors in parentheses

Parameter	Coefficient
M_0	0.9506 (0.0091)
V_{max}	3.0085 (0.9869)
b_{lo} †	0.00
b_{hi} †	1.00
Shape Parameters for V_i ††	
a	1.00
b	0.75
Shape Parameters for b_i ††	
a	0.50
b	1.25
log$L$	-170.69

†Coefficient constrained to equal the indicated value.
††Coefficient obtained via grid search.

Table 5. Profit and Welfare Consequences of Versioning

	Optimal Versioning	Single Price Monopoly	Welfare Maximum
N_1	2	NA	NA
P_1	\$1.28	NA	NA
N_2	12	12	12
P_2	\$4.84	\$4.30	\$0.00
% Purchasing Version 1†	24.9%	NA	NA
% Purchasing Version 2†	26.1%	33.0%	100.0%
Profit per Customer†	\$1.58	\$1.42	\$0.00
Consumer Surplus per Customer†	\$1.29	\$1.36	\$3.99
Total Surplus per Customer†	\$2.87	\$2.78	\$3.99

†Averaged only over customers with non-zero valuations.

Figure 1. Optimal Versioning

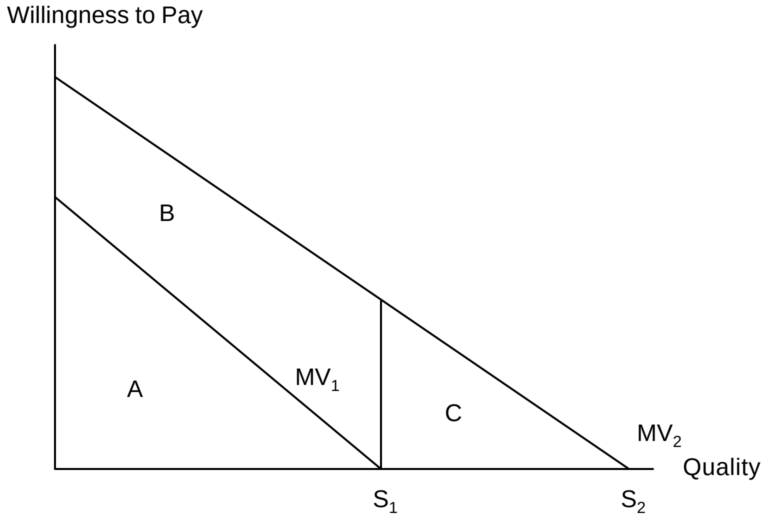


Figure 2. Optimal Versioning

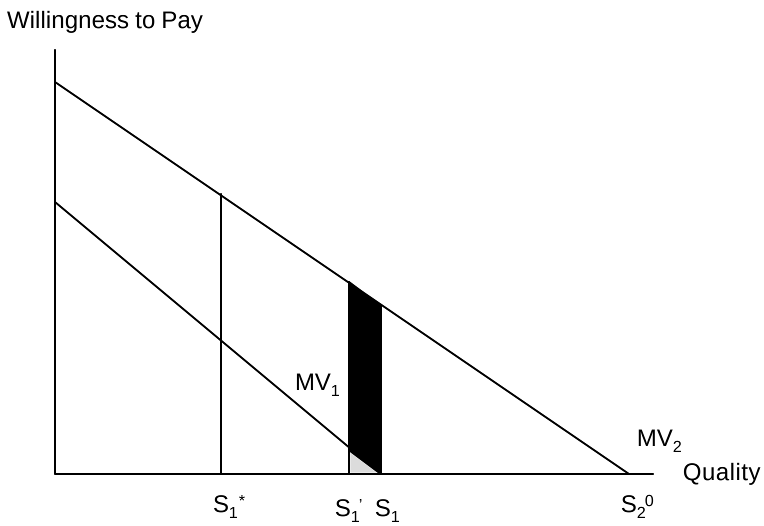
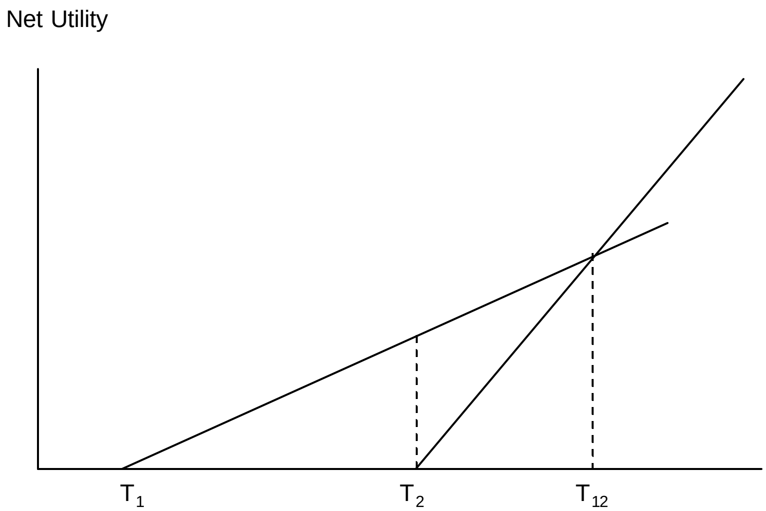


Figure 3. Net Utility for Two Versions of a Good



Appendix

Sample Images Included in Sale Bundles

(Some images were animated)