Versioning in the Information Economy: Theory and Applications

Paul Belleflamme

Abstract: Price discrimination consists in selling the same product to different buyers at different prices. When sellers cannot relate a buyer's willingness to pay to some observable characteristics, price discrimination can be achieved by targeting a specific package (i.e., a selling contract that includes various clauses in addition to price) for each class of buyers. The seller faces then the problem of designing the menu of packages in such a way that each consumer indeed chooses the package targeted for her. This practice, known as versioning (or as second-degree price discrimination), is widespread in the information economy. In this paper, we propose a simple unified framework to study the general theory behind versioning and to consider a number of specific versioning strategies used in the information economy (namely, bundling, functional degradation and conditioning prices on purchase history). (JEL L82, L86, K11, O34)

1 Introduction

The expression “information economy” (or the alternative buzz words “new”, “cyber”, “network” and “e- economy”) is used with considerable – and sometimes excessive – frequency nowadays. To give the expression proper contents, let us make clear that the general idea of an “information economy” includes two interrelated notions. On the one hand, the information economy refers to the industries primarily producing, processing, and distributing information; these industries form together a so-called “information sector” which contributes an increasing share to wealth and job creation. On the other hand, the information economy also encompasses the idea that every industry makes an increasing use of information and information technology to reorganize, make themselves more productive, and create new ways of doing business.
To clarify matters further, it is convenient to divide the objects of economic transactions in the information economy into two complementary sides: the content (or information) side and the infrastructure (or technology) side. On the content side, the basic unit that is transacted is information. Following Varian (1998), we take information very broadly as anything that can be digitized (i.e., encoded as a stream of bits): text, images, voice, data, audio and video. Basic information is transacted under a wide range of formats or packages (which are not necessarily digital). These formats are generically called information goods. Books, movies, music, magazines, software, games, databases, telephone conversations, stock quotes, web pages, news, ring tones, etc all fall into this category. On the infrastructure side, we define as information technologies all the technologies for recording, conditioning, transmitting, distributing, using and processing information. Examples of information technologies are hardware and software platforms for office automation systems, telecommunication equipment such as servers, bridges, routers, hubs and wiring.

Information goods have the distinguishing characteristic of involving high fixed costs but low (often zero) marginal costs. In this case, cost-based pricing is not a sensible approach for firms and must give way to value-based pricing: an information good must be priced according to the value consumers attach to it, not according to its production cost. Moreover, as different consumers generally attach very different values to the same information good, the producer should set not a single but several value-based prices for its information good. This practice is known as price discrimination. More precisely, price discrimination implies that two varieties of a good are sold (by the same seller) to two buyers at different net prices, the net price being the price (paid by the buyer) corrected for the cost associated with the product differentiation.3

We begin by addressing the feasibility of price discrimination. In a perfectly competitive market, there cannot be two different prices for the same product, for it would otherwise be possible to earn a profit by engaging in arbitrage, i.e., by buying at the low price and selling at the high price. Implicit behind this “law of one price” are the assumptions that arbitrage is perfect (costless) and that agents are perfectly informed about the different prices. However, in the real world (and especially in the information economy), examples abound in which different prices are observed for what is apparently the same product.

---

3 This definition is adapted from Philips (1983, 6). As Philips argues, this definition seems more acceptable than the standard definition, which identifies price discrimination to the practice of setting different prices for the same product, but differentiated products, that are sold at discriminatory prices (see the practice of versioning below).
First, it must be the case that firms enjoy some market power, so that they are in a position to set prices. Second, letting aside the possibility of imperfect consumer information, it must be that arbitrage costs are so high that consumers do not find it profitable to transfer the good between them.\(^4\)

Let us now distinguish between different types of price discrimination. Following Pigou (1920), it is customary to distinguish three different types, according to the information that firms have about buyers. The most favorable case for the firm is when it has complete information about individual preferences. The firm is then able to charge an individualized price for each buyer and for each unit purchased by each buyer, thereby extracting all of the consumer surplus. In Pigou’s taxonomy, this practice is known as first-degree price discrimination (or perfect discrimination). Shapiro and Varian (1999) propose the more descriptive term of personalized pricing. It is often argued that personalized pricing cannot be applied in practice because of the enormous amount of information it requires. However, firms are now able to use information technologies in order to improve their knowledge of consumers’ preferences and, thereby, to personalize price offers.\(^5\)

When the firm does not know exactly each consumer’s willingness to pay, it may still manage to extract a fraction of the consumer surplus by relying on some indicators (such as age, occupation, location) that are related to the consumers’ preferences. If the firm can observe a buyer’s characteristics, then it can charge different prices as a function of these characteristics. This type of price discrimination is referred to as third-degree price discrimination in Pigou’s taxonomy, or as group pricing in Shapiro and Varian’s. As group pricing can be seen as a special case of a multiproduct firm’s pricing problem (see Tirole 1988), we shall not discuss it further in this paper.

When buyers’ characteristics are not directly observable, the firm still has the option to use self-selecting devices in order to extract some consumer surplus. The idea is to discriminate between heterogeneous buyers by targeting a specific package (i.e., a selling contract that includes various clauses in addition to price) for each class of buyers. The firm faces then the problem of designing the menu of packages in such a way that each consumer indeed chooses the package targeted for her. Pigou describes this practice as second-degree price discrimination, whereas Shapiro and Varian prefer to refer to it, more descriptively, as versioning. This practice is our main focus of attention

\(^4\) As Tirole (1988) points out, arbitrage costs might be associated either with the transferability of the good itself (e.g., it is too time-consuming to unbundle a “3-for-the-price-of-2 package” in order to re-sell individual units), or with the transferability of demand between different packages aimed at different consumers (see Section 2).

\(^5\) See, e.g., Ulph and Vulkan (2000) for a study of personalized pricing in a duopoly setting.
in this paper. In Section 2, we present an integrated model which allows us to study how to implement versioning and when it is optimal to do so. Then, in Section 3, we consider a number of specific strategies for discriminating between consumers in the information economy: bundling or tying different goods together, creating a new version by disabling some functions of an existing version, and conditioning prices on purchase history. Section 4 gives some concluding remarks.

2 Versioning: how and when?

In the nineteenth century, Dupuit (a French engineer and economist, quoted by Ekelund 1970) analyses the practice of the three-class rail system as follows:

“It is not because of the few thousand francs which would have to be spent to put a roof over the third-class carriages or to upholster the third-class seats that some company or other has open carriages with wooden benches... What the company is trying to do is prevent the passengers who can pay the second-class fare from travelling third-class; it hits the poor, not because it wants to hurt them, but to frighten the rich... And it is again for the same reason that the companies, having proved almost cruel to third-class passengers and mean to second-class ones, become lavish in dealing with first-class passengers. Having refused the poor what is necessary, they give the rich what is superfluous.”

What Dupuit describes is a classical example of versioning: by offering the same product under a number of “packages” (i.e., some combinations of price and product characteristics), the seller is able to sort consumers according to their willingness to pay. The key is to identify some dimensions of the product that are valued differently across consumers, and to design the product line so as to emphasize differences along those dimensions. The next step consists in pricing the different versions in such a way that consumers will sort themselves out by selecting the version that most appeals to them.

Examples of such practice abound in the information economy. The dimension along which information goods are versioned is usually their quality, which is to be understood in a broad sense (for instance, the quality of a software might be measured by its convenience, its flexibility of use, the performance of the user interface, ...). Versioning of information goods can also be based on time or on quantity.

---

6 Shapiro and Varian (1998) give a series of specific examples. For instance, referring to “nagware” (i.e., a form of shareware that is distributed freely but displays a screen encouraging
We now develop a simple model to understand how a monopolist should choose prices to induce self-selection of the consumers between different versions of the product. We also identify conditions under which versioning allows the monopolist to increase profits.

2.1 A simple model of versioning

We consider the problem of a monopolist choosing packages of quality and price level for an information good when consumers have unit demands. The monopolist has identified one particular dimension of the good for which consumers have different value. In particular, we suppose that there is a continuum of potential consumers who are characterized by their valuation, \( \theta \), for this dimension of the information good. We assume that the "taste parameter" \( \theta \) is uniformly distributed on the interval \([0,1]\). The monopolist can produce this dimension of the good at two levels of quality, which are given exogenously. We note the two levels \( s_1 \) and \( s_2 \), with \( s_2 > s_1 \). As for the other dimensions of the information good, consumers are assumed to share the same valuation. Consumers’ preferences are then described as follows: when consuming a unit of the good of quality \( s_i \) sold at price \( p_i \), a consumer of type \( \theta \) enjoys a (net) utility of

\[
U(\theta, s_i) = k + \theta - p_i,
\]

where \( k \geq 0 \) represents the common valuation for the other dimensions of the information good. We assume that \( k < s_1 \), meaning that the consumer with the highest taste parameter (\( \theta = 1 \)) values the particular dimension (at its lowest quality level) more than all the other dimensions of the good. Finally, we pose that a consumer’s utility is zero if she refrains from buying.
The monopolist faces the following problem. He knows the aggregate distribution of the taste parameters but is unable to identify a particular consumer's type. Two options are available: he can either sell a unique quality at a single price, or price-discriminate by offering the two qualities at different prices. Let us examine the two options in turn, assuming that there is a constant marginal cost \( c_i \geq 0 \) to produce one unit of the good of quality \( s_i \) (with \( c_i \leq k \) for \( i = 1, 2 \), meaning that the less eager consumer would buy either version if priced at marginal cost). Assumption 1 summarizes the relationship between the parameters of the model. For reasons that will become clear immediately, we also assume that the cost of producing the high quality \( (c_2) \) is not too large with respect to the cost of producing the low quality \( (c_1) \). Assumption 2 makes this condition precise.

**Assumption 1** \( 0 \leq c_1, c_2 \leq k < s_1 < s_2 \).

**Assumption 2** \( c_2 < c_1 + x \), with \( x = k + s_2 - c_1 - \sqrt{s_2^2 / s_1 (k + s_1 - c_1)} > 0 \).

### Selling a single quality

Under Assumption 2, the monopolist will choose to produce quality \( s_2 \) if he decides to sell a single quality. For a given price \( p_2 \) of quality \( s_2 \), the marginal consumer who is indifferent between buying and not buying is identified by \( \theta_{20}(p_2) \), which solves \( k + \theta_{20}(p_2) s_2 - p_2 = 0 \Leftrightarrow \theta_{20}(p_2) = (p_2 - k) / s_2 \). Since all consumers with a higher valuation than \( \theta_{20}(p_2) \) will buy the good, the monopolist's profit-maximization problem writes as

\[
\max_{p_2} \pi_{1q} = (p_2 - c_2) \left[ 1 - \theta_{20}(p_2) \right] = (p_2 - c_2) \left( 1 - \frac{p_2 - k}{s_2} \right).
\]

The optimal price and profit are easily computed as

\[
\hat{p}_2 = \frac{k + s_2 + c_2}{2} \quad \text{and} \quad \hat{\pi}_{1q} = \frac{(k + s_2 - c_2)^2}{4s_2}.
\]

### Selling the two qualities

We need now to find the profit maximising price pair \((p_1, p_2)\) that induces some consumers to select quality \( s_1 \) and other consumers to select quality \( s_2 \). Given \((p_1, p_2)\), we denote by \( \theta_{12} \) the marginal consumer who is indifferent
between consuming either of the two qualities, and by $\theta_{10}$, the marginal consumer who is indifferent between consuming quality $s_1$ and not consuming at all. By definition, we have

\[ k + \theta_{12}s_1 - p_1 = k + \theta_{12}s_2 - p_2 \Leftrightarrow \theta_{12}(p_1, p_2) = \frac{p_2 - p_1}{s_2 - s_1}, \]

\[ k + \theta_{10}s_1 - p_1 = 0 \Leftrightarrow \theta_{10}(p_1) = \frac{p_1 - k}{s_1}. \]

To achieve price discrimination, prices must be such that $0 \leq \theta_{10} < \theta_{12} < 1$. In that case, the population of consumers is divided into two (possibly three) groups: consumers with $\theta \geq \theta_{12}$ buy quality $s_2$, while consumers with $\theta_{10} \leq \theta < \theta_{12}$ buy quality $s_1$ (and if $\theta_{10} > 0$, consumers with $0 \leq \theta \leq \theta_{10}$ do not buy any quality). Prices must thus satisfy the following two constraints.

(A) \[ \theta_{12}(p_1, p_2) < 1 \Leftrightarrow p_2 < p_1 + (s_2 - s_1). \]

(B) \[ \theta_{10}(p_1) < \theta_{12}(p_1, p_2) \Leftrightarrow s_1 / (p_1 - k) > s_2 / (p_2 - k). \]

These are the so-called self-selection constraints (also known as “incentive-compatibility constraints”). If the menu $(s_i, p_i)_{i=1,2}$ is to be feasible in the sense that it will be voluntarily chosen by the consumers, then consumers of each group must prefer consuming the package intended for them as compared to consuming the other group’s package or not consuming any package. Constraint (A) states that the price of the high-quality version must be lower than the price of the low-quality version, augmented by the quality gap (as valued by the most eager consumer, i.e., $\theta = 1$). This is a necessary condition for positive sales of the high-quality version. Constraint (B) states that the low-quality version must offer a better “quality-price ratio” (computed here as $s_i / (p_i - k)$ for version $i$) than the high-quality one. This is a necessary condition for positive sales of the low-quality version.

Expressing $p_2$ as $p_1 + \Delta$, we can write the monopolist’s profit-maximization problem as follows

\[
\max_{p_1, \Delta} \pi_{2q} = (p_1 - c_1)\left[\theta_{12}(\Delta) - \theta_{10}(p_1)\right] + (p_1 + \Delta - c_2)\left[1 - \theta_{12}(\Delta)\right]
\]

s.t. (A) and (B) are met.

That is, the monopolist sets the price of the low-quality version $(p_1)$ and the premium over that price associated with the high-quality version $(\Delta)$. The two first-order conditions are given by
The former condition shows that an increase in \( p_1 \) has a twofold effect on profits: on the one hand, revenues are gained on the consumers of the two versions but on the other hand, margins are lost from consumers of the low-quality version who leave the market because of the price increase. The latter condition indicates two similar effects: a higher premium \( \Delta \) gives rise to increased revenue from the consumers of the high-quality version but makes some consumers switch to the low-quality version (which is sold at a margin \( p_1 - c_1 \) instead of \( p_2 - c_2 \)).

Replacing \( \theta_{12} \) and \( \theta_{10} \) by their respective value and solving for each first-order condition, we compute the profit-maximizing prices as

\[
p^*_1 = \frac{k + s_1 + c_1}{2}, \quad \Delta^* = \frac{c_2 - c_1 + s_2 - s_1}{2} \Rightarrow p^*_2 = \frac{k + s_2 + c_2}{2}.
\]

Do these prices meet the constraints? We first check that the market is not fully covered at these prices: our assumption that \( s_1 > k \) implies that \( \theta_{10}(p^*_1) > 0 \), so that consumers with very low values of \( \theta \) do not purchase any version. Next, a few lines of computations establish for which regions of parameters the self-selection constraints are met:

(A') \[ p^*_2 < p^*_1 + (s_2 - s_1) \Leftrightarrow c_2 < c_1 + (s_2 - s_1) \]

(B') \[ s_1 / (p^*_1 - k) > s_2 / (p^*_2 - k) \Leftrightarrow s_1 / (c_1 - k) > s_2 / (c_2 - k) \]

Conditions (A') and (B') are nothing but the self-selection constraints (A) and (B) expressed in the case of marginal-cost pricing (i.e., \( p_1 = c_1 \) and \( p_2 = c_2 \)).

We now discuss the optimality of versioning under different scenarios for the values of the parameters \( k, c_1, \) and \( c_2 \).

2.2 When is versioning optimal?

We have just shown that conditions (A') and (B') are necessary for versioning to be feasible. Using a simple “revealed preference” argument, we can also say that versioning is more profitable than selling a single version when conditions (A') and (B') are met. Indeed, if the monopoly decides to set a pair of prices that induces some consumers to purchase the low-quality version, it is because this strategy
increases profits (otherwise, he would choose prices such that only the high-quality version is purchased).

**Cannibalization vs market expansion**

To understand the importance of the two conditions, let us detail the two conflicting effects versioning induces on the monopolist's profits (see Figure 1). Under conditions \((A')\) and \((B')\), the marginal consumers in the two options are ranked as follows:

\[
0 < \theta_{10}(p_1^*) < \theta_{20}(\hat{p}_2) < \theta_{12}(p_1^*, p_2^*) < 1.
\]

Figure 1

**Effects of versioning on firm’s profits**

That is, the effect of versioning on consumers’ choices is twofold: first, because \(\theta_{20}(\hat{p}_2) < \theta_{12}(p_1^*, p_2^*)\), fewer consumers buy the high-quality good; second, because \(\theta_{10}(p_1^*) < \theta_{20}(\hat{p}_2)\), some previous non-consumers now buy the low-quality good. How does this affect profits? The first thing to note is that nothing changes for the consumers with a very high taste parameters \(\theta\), in the two options, they buy quality \(s_2\) at the same price \((p_2^* = \hat{p}_2)\). Going down the distribution of \(\theta\), we encounter consumers who would buy quality \(s_2\) if it were the only quality available, but who would buy quality \(s_1\) otherwise. This cannibalization effect has the following negative impact on the monopolist’s profit:

\[
d\pi_{ca} = \left[ \theta_{12}(p_1^*, p_2^*) - \theta_{20}(\hat{p}_2) \right] \left[ (p_1^* - c_1) - (\hat{p}_2 - c_2) \right] \left[ s_2 - s_1 - (c_2 - c_1) \right] / 4s_2(s_2 - s_1) < 0.
\]
Finally, versioning has a market expansion effect: there are consumers who would buy quality $s_1$ when it is offered but who would not buy quality $s_2$ were it the only quality available. The positive impact on profit is equal to

$$d\pi_{me} = \left[ \theta_{20}(p_2) - \theta_{10}^*(p_1^*) \right] \left( p_1^* - c_1 \right) = \frac{\left[ s_2(k-c_1) - s_1(k-c_2) \right] \left( k + s_1 - c_1 \right)}{4s_1s_2} > 0.$$ 

Adding the two effects, we find that under conditions (A') and (B'), price discrimination is indeed the most profitable option:

$$\pi_2(q^*(p_1^*,p_2^*)) - \pi_1(q(p_2)) = d\pi_{ca} + d\pi_{me} = \frac{\left[ s_2(k-c_1) - s_1(k-c_2) \right]^2}{4s_1s_2(s_2 - s_1)} > 0.$$

**Is versioning optimal for information goods?**

Information goods have the distinguishing characteristic of involving high fixed costs but low (often zero) marginal costs. More generally, the marginal cost of production is invariant with product quality. In our setting, this would mean that $c_1 = c_2 = c$. With $c$ being near zero, we observe that assuming $0 < k < s_1$ suffices to guarantee that conditions (A') and (B') are met and that versioning is optimal.\(^9\) Note that for $c_1 = c_2 = 0$, we have that

$$d\pi_{ca} + d\pi_{me} = -\frac{s_2 - s_1}{4s_2} k + \frac{(s_2 - s_1)(k + s_1)}{4s_1s_2} k = \frac{s_2 - s_1}{4s_1s_2} k^2 > 0 \Leftrightarrow k > 0.$$

We record our main result below.

Suppose the consumers’ utility for the information good can be separated along two dimensions: a “key dimension” for which consumers have different valuations, and a “secondary dimension” for which all consumers have the same, positive, valuation. Suppose also that some consumers value the key dimension more than the secondary dimension, and that the marginal cost of producing any level of quality for the key dimension is near zero. Then versioning the information good along the key dimension is the most profitable option for the monopolist.

\(^9\) However, if $k = 0$ and $c_1 = c_2$, condition (B') cannot be met and the monopolist will prefer to offer the high quality only. This is the result reached by Salant (1989) under the assumption that the marginal cost function for quality is linear.
**Damaged goods**

One extreme form of versioning occurs when firms intentionally damage a portion of their goods in order to price discriminate. Such “damaged good strategy” is widely used in software markets: initially, the producer develops a complete full-featured version and then introduces additional low-quality versions by degrading quality of the original version. Denekere and McAfee (1996) report other instances where firms actually incur an extra cost in order to produce the low-quality version. They model this extra cost by assuming that the marginal cost of production is higher for the low quality version. In the above setting, this would mean that \( c_1 > c_2 \geq 0 \). Continuing to assume that \( 0 < k < s_1 \), let us examine whether versioning could be optimal under this alternative scenario. Condition (A') is clearly satisfied as \( c_2 - c_1 < 0 \). As for condition (B'), it can be satisfied if the “damaging cost” \( c_1 - c_2 \) is not too large. We conclude that versioning can be feasible, and thus optimal, even if it is more costly to produce the low-quality version of the product.

Looking at the examples provided in Denekere and McAfee (1996), one could argue that the damaged good strategy is more likely to require some additional fixed cost rather than an increase in marginal cost. In this case, supposing \( c_1 = c_2 = 0 \) and letting \( F > 0 \) denote the fixed cost of creating quality \( s_1 \) by damaging quality \( s_2 \), we have that the damaged good strategy is optimal if and only if

\[
\pi_{2d}(\hat{p}_1, \hat{p}_2) - F > \pi_1(\hat{p}_2) \iff F < \frac{s_2 - s_1}{4s_1s_2}k^2.
\]

### 3 Versioning: Applications

We discuss now three specific ways, observed in the information economy, of inducing consumers’ self-selection in order to capture a larger share of the consumer surplus: bundling, functional degradation, and conditioning prices on purchase history.

---

10 For instance, *Sony recordable MiniDiscs* (MDs) come in two formats (60-minute and 74-minute discs), which are sold at different prices. Yet, the two formats are physically identical: a code in the table of contents identifies a 60-minute disc and prevents recording beyond this length, even though there is room on the media.
3.1 Bundling

Just as inducing self-selection by offering a menu of versions enhances the monopolist’s ability to extract surplus, so can selling different products as a combination package. Two such techniques are bundling and tying. The practice of bundling consists in selling two or more products in a single package (bundling is said to be “pure” when only the package is available, or “mixed” when the products are also available separately). The distinguishing feature of bundling is that the bundled goods are always combined in fixed proportions. In contrast, the related practice of tying (or tie-in sales) is less restrictive in that proportions might vary in the mix of goods. It takes only a little reflection to recognize how common practices bundling and tying are in the information economy. Examples abound both on the content side and on the infrastructure side, as illustrated below.

**Examples of bundling in the information economy**

*Content side.* (i) Subscription to cable television is typically to a package of channels together, rather than to each channel separately; similarly for subscription to magazines or for CDs (which can be seen as bundles of different songs). (ii) Software companies sell individual products but also offer packages (or “suites”) consisting of several applications (e.g., Microsoft Office suite). (iii) Movie distributors frequently force theaters to acquire “bad” movies if they want to show “good” movies from the same distributor.

*Infrastructure side.* (i) Computer manufacturers offer bundles that include both a computer (a central processing unit) and a monitor. (ii) Audio equipment usually can be bought as separate components or as a complete system. (iii) Photocopier manufacturers offer bundles that include the copier itself as well as maintenance; they also offer the alternative of buying the copier and servicing separately. (iv) A classic example of tying was the practice adopted by IBM in the era of punch-card computers: IBM sold its machines with the condition that the buyer use only IBM-produced tabulating cards. Current examples involve computer printers (ink cartridges are generally specific to a particular model of a particular manufacturer) or some photographic films (only Polaroid films fit a Polaroid Instamatic).

Economists have given different explanations for bundling and tying. First, some explanations are too transparent to merit formal treatment. In the case of perfectly complementary products, such as matching right and left shoes, no
one questions the rationale of bundling: there is virtually no demand for separate products and bundling them together presumably conserves on packaging and inventory costs. In other cases where products are not necessarily complements, various cost efficiencies might provide a basis for profitable bundling. More interestingly, even in the absence of cost efficiencies, there are demand side incentives that makes bundling and tying profitable strategies. On the one hand, both practices can serve as an effective tool for sorting consumers and price discriminate between them; it is this rationale we concentrate on in this section.11 On the other hand, bundling and tying are also particularly effective entry-deterrent strategies; the recent case brought against Microsoft by the European Commission follows this line of argument.12

We now present a simple example to illustrate the use of bundling as an alternative strategy for (second-degree) price discrimination.13 Consider two monopolized products, which are independent both in terms of demand (i.e., the value a consumer places on one product does not depend on the consumption of the other product) and in terms of costs (i.e., there are no cost advantages of multiproduct production; in particular, we assume zero marginal cost for both goods). We still have a continuum of potential consumers who are characterized by a taste parameter $\theta$, which is assumed to be uniformly distributed over the interval $[0,1]$. Consumers have a unit demand for each of two information goods (indexed by $i = 1, 2$). A typical consumer $\theta$ has the following (gross) utility function:

$$
\begin{align*}
    u_1(\theta) &= \theta & \text{if consuming good 1,} \\
    u_2(\theta) &= a\theta + (1-\alpha)(1-\theta) & \text{if consuming good 2,} \\
    u_b(\theta) &= u_1(\theta) + u_2(\theta) & \text{if consuming the bundle.}
\end{align*}
$$

where $\alpha \in [0,1]$ measures the correlation between the two distributions of utilities across consumers: for $\alpha$ in $[0,1/2]$ there is perfectly negative correlation, for $\alpha = 1/2$ there is no correlation, and for $\alpha$ in $(1/2,1]$ there is perfectly positive correlation. Note that for $\alpha = 1/2$ the valuation for product 2 is constant across consumers.

---

11 For related literature, see Adams and Yellen (1976), Schmalensee (1984) or McAfee, McMillan and Whinston (1989).
12 Three papers addressing this topic are Whinston (1990), Choi and Stefanidis (2003) and Nalebuff (2004).
13 Regarding tying, let us simply mention that it can be viewed as a price discrimination device because it enables the monopolist to charge more to consumers who value the good the most. Here, the value consumers place on the primary product (e.g., the printer) depends on the frequency with which they use it, which is itself measured by their consumption of the tied product (e.g., the ink cartridges). Those who most need the primary product will consume more of the secondary product and, thereby, pay a higher effective price.
We now compare two options for the monopolist: either selling the two goods separately (separate sales), or selling them together as a bundle (pure bundling).

**Separate sales**

Let $p_i$ denote the price of good $i$ ($i = 1, 2$). We identify two pivotal consumers.

- Consumer $\theta_1(p_1)$ is such that $u_1(\theta) - p_1 = 0$: this consumer is indifferent between buying good 1 only and not buying any good (or between buying the two goods and buying good 2 only); as $u_1(\theta)$ is an increasing function of $\theta$, consumers with a value of $\theta$ larger than $\theta_1(p_1)$ strictly prefer the first option.

- Consumer $\theta_2(p_2)$ is such that $u_2(\theta) - p_2 = 0$: this consumer is indifferent between buying good 2 only and not buying any good (or between buying the two goods and buying good 1 only); if $\alpha \geq 1/2$ (positive correlation), $u_2(\theta)$ is an increasing function of $\theta$ and consumers with a value of $\theta$ larger than $\theta_2(p_2)$ strictly prefer the first option; otherwise (negative correlation), $u_2(\theta)$ is a decreasing function of $\theta$ and the consumers who strictly prefer the first option are those with a value of $\theta$ smaller than $\theta_2(p_2)$.

Consider first the case of negative correlation ($0 \leq \alpha < 1/2$). Using the definition of the pivotal consumers, we see that a consumer would prefer to buy nothing if her value of $\theta$ was comprised between $\theta_2(p_2)$ and $\theta_1(p_1)$: as $\theta > \theta_2(p_2)$, she is better off buying nothing than buying good 2 only; moreover, as $\theta < \theta_1(p_1)$, she is also better off buying nothing than buying good 1 only. Inverting the argument, we conclude that the monopolist will cover the whole market (i.e., sell at least one good to each and every consumer) if he sets prices so that

$$\theta_1(p_1) = p_1 \leq \theta_2(p_2) = \frac{1 - \alpha - p_2}{1 - 2\alpha}.$$

Under this condition, consumers are split into three groups: those with $0 \leq \theta < \theta_1(p_1)$ buy good 2 only, those with $\theta_1(p_1) \leq \theta < \theta_2(p_2)$ buy the two goods, and those with $\theta_2(p_2) \leq \theta < 1$ buy good 1 only. The monopolist’s maximization program writes thus as
The unconstrained prices are easily found $p_1^* = 1/2 > 0$ as and $p_2^* = (1 - \alpha) / 2$. One checks that $(1 - 2\alpha)p_1^* < 1 - \alpha - p_2^*$. However, the last constraint is satisfied if and only if $\alpha \leq 1/3$. For $1/3 < \alpha < 1/2$, we have a corner solution: $\bar{p}_2 = \alpha$ and no consumer buys good 1 only. A quick analysis reveals that the latter solution also holds in the special case where $\alpha = 1/2$ and all consumers have the same utility for good 2.

Consider next the case of positive correlation $(1/2 < \alpha < 1)$. Now, by the definition of the pivotal consumers, we find that consumers with a value of $\theta$ larger than $\max \{\theta_1(p_1), \theta_2(p_2)\}$ buy both goods, while those with a value of $\theta$ comprised between $\min \{\theta_1(p_1), \theta_2(p_2)\}$ and $\max \{\theta_1(p_1), \theta_2(p_2)\}$ buy only a single good. Whatever the ranking of $\theta_1(p_1)$ and $\theta_2(p_2)$, the monopolist’s profit writes now as

$$\pi_s = p_1(1-p_1) + p_2 \left(1 - \frac{p_2 - (1-\alpha)}{2\alpha - 1}\right).$$

The profit-maximising prices are: $p_1^* = 1/2$ and $p_2^* = \alpha / 2$. At these prices, one checks that $\theta_2(p_2^*) < \theta_1(p_1^*) < 1$. For the solution to be interior, we still need that $\theta_2(p_2^*) > 0 \Leftrightarrow \alpha > 2/3$. If the latter condition is not satisfied, we have a corner solution: $\bar{p}_2 = 1 - \alpha$, and the market is fully covered.

Collecting our previous results, we can compute the optimal profit under separate sales for all values of $\alpha$:

$$\pi_s^* = \begin{cases} 
\frac{\alpha^2 - 4\alpha + 2}{4(1-2\alpha)} & \text{for } 0 \leq \alpha \leq \frac{1}{3}, \\
\alpha + \frac{1}{4} & \text{for } \frac{1}{3} < \alpha \leq \frac{1}{2}, \\
1 - \alpha + \frac{1}{4} & \text{for } \frac{1}{2} < \alpha \leq \frac{2}{3}, \\
\frac{\alpha^2 + 2\alpha - 1}{4(2\alpha - 1)} & \text{for } \frac{2}{3} < \alpha \leq 1.
\end{cases}$$
Pure bundling

Let $p_b$ denote the price of the bundle (which is the only available good in the present case). Consumer $\theta$’s net utility is given by $u_b(\theta) - p_b = (1 - \alpha) + 2\alpha \theta - p_b$. Consider first the special case of perfect negative correlation ($\alpha = 0$). In that case, all consumers have the same utility for the bundle: $u_b(\theta) = 1 - \alpha = 1$. The monopolist will thus set $p_b = 1$ and sell the bundle to the whole population of consumers, achieving a profit of $\pi_b = 1$.

For $\alpha > 0$, we can identify the consumer who is indifferent between buying the bundle or not as $\theta_* = (p_b - 1 + \alpha) / 2\alpha$. The monopolist’s profit-maximization program is given by

$$\max_{p_b} \pi_b = p_b \left(1 - \frac{p_b - 1 + \alpha}{2\alpha}\right) \text{ s.t. } 1 - \alpha \leq p_b \leq 1 + \alpha.$$

Solving for the first-order condition, we find $p_b^* = (1 + \alpha) / 2$. This solution meets the constraints if and only if $\alpha \geq 1 / 3$. Otherwise, we have a corner solution: $p_b^* = 1 - \alpha$ and the market is fully covered. In sum, the optimal profit under pure bundling is equal to

$$\pi_b^* = \begin{cases} 1 - \alpha & \text{for } 0 \leq \alpha < \frac{1}{3} \\ \frac{(1 + \alpha)^2}{8\alpha} & \text{for } \frac{1}{3} \leq \alpha < 1. \end{cases}$$

When is pure bundling more profitable than separate sales?

Comparing expressions (3) and (5) for all values of $\alpha$, we find the following (see Figure 2):

- for $0 \leq \alpha \leq \frac{1}{3}$, $\pi_b^* = 1 - \alpha > \pi_s^* = \frac{\alpha^2 - 4\alpha + 2}{4(1 - 2\alpha)}$;
- for $\frac{1}{3} < \alpha < \sqrt{\frac{1}{7}}$, $\pi_b^* = \frac{(1 + \alpha)^2}{8\alpha} > \pi_s^* = \alpha + \frac{1}{4}$;
- for $\sqrt{\frac{1}{7}} < \alpha < \frac{1}{2}$, $\pi_b^* = \frac{(1 + \alpha)^2}{8\alpha} < \pi_s^* = \alpha + \frac{1}{4}$.
We therefore conclude

Profits are higher under bundling than under separate sales if and only if the correlation between the distributions of consumer utilities for the two goods is sufficiently negative i.e., if and only if \( \alpha \leq \sqrt{1/7} \approx 0.38 \).

**Figure 2**

**Comparison of profits under separate sales and pure bundling**

The intuition for this result is simple. By selling a bundle at a lower price, the monopolist attempts to attract consumers who place a relatively low value on either of the two goods but who are willing to pay a reasonable sum for the bundle. When the two goods are sold separately, these consumers would buy a single good, but when the goods are sold as a bundle, they would buy the bundle and, therefore, acquire a good they would not have purchased other-
wise. Naturally, for such consumers to exist, the variation in consumer valuations of the goods must be significant.

### 3.2 Functional degradation

We now examine the logic behind a practice that becomes increasingly common nowadays: the *functional degradation of computer software*. Well-known examples of this practice are software like *Acrobat Reader*, *RealPlayer*, various *Microsoft Office* (*Word*, *Excel*, *PowerPoint*) *Viewers*. These software are designed to view and print (or play) contents written in a specific format, but are not capable of producing the contents in the specific format. The software manufacturers provide typically two different versions of a product, the read-only (or play-only) version and the full version. They offer the viewer or player almost free of charge by allowing consumers to download it on the Internet. However, the viewer (or player) is viewing and printing (or playing) only. To be able to create and edit contents (as well as viewing or playing them), users need to purchase the corresponding full version, which is of course sold at a positive price. In the following simple model, based on Csorba and Hahn (2003), we show that this practice is based on a mixture of economic motivations: versioning, damaged good strategy, network building through free versions, and unbundling.

Suppose a firm is the sole supplier of a software which combines a read and a write function: the write function is required to produce documents that can be read using the read function. Since not all consumers are interested in writing documents, it is natural to offer the software under two versions: a *full* (write and read) version and a *read-only* version.\(^{14}\) As before, we assume that there is a continuum of potential consumers who buy at most one unit of the good. A consumer is identified by a parameter $\theta$ (drawn from a uniform distribution on the unit interval), which indicates her valuation for the two functions of the good. More precisely, the valuations for the two functions are assumed to be proportional: $\theta$ is the valuation for the write function, while $\beta \theta$ (with $\beta > 0$) is the valuation for the read function. That is, we restrict the attention to cases where users who value highly one function also value highly the other function (and vice versa).\(^{15}\)

\(^{14}\) A “write-only” version does not make much sense, as it seems hard to write a document without being able to read it.

\(^{15}\) In other words, there is positive correlation between the distribution of utilities for the two function. Applying the analysis of the previous section, we can anticipate that this positive correlation is likely to drive the monopolist to “unbundle” the two functions.
The software exhibits two-sided network effects, insofar as the users’ utility from reading (resp. writing) increases with the number of writers (resp. readers). Letting \( p_f \) and \( p_r \) denote the prices of the full and of the read-only version, and \( n_f \) and \( n_r \) denote the number of consumers who buy the full and the read-only versions, we can express the (net) utility of consumer \( \theta \) as follows:

\[
U(\theta) = \begin{cases} 
\theta(n_f + n_r) + \beta \theta n_f - p_f & \text{if buying the full version,} \\
\beta \theta n_f - p_r & \text{if buying the read-only version,} \\
0 & \text{if not buying.}
\end{cases}
\]

Let us detail this utility function. The full version can be seen as a bundle combining the two functions. Utility is assumed to be additively separable in the two functions. The first term is the utility from the write function: a user owning the full version enjoys a network effect exerted by the users able to read the documents she produces (i.e., those users who own either version of the good). The second term is the utility from the read function: the network effect, here, is exerted by the users able to write documents and is valued at \( \beta \theta \) (instead of \( \theta \) for the write function). Naturally, utility from owning the read-only version is limited to the latter term.

Because the firm cannot observe the consumers’ types, it cannot resort to personalized pricing or group pricing. Two options are available: either sell the full version only (which can be seen as pure bundling), or sell the two versions (which is a form of versioning or of mixed bundling). We consider the two options in turn.

**Sell the full version only**

When selling only the full version, the firm is in the position of a monopolist pricing a good that exhibits network effects. If the read-only version is not available, \( n_r \) is necessarily equal to zero and the consumer \( \theta \)'s options are restricted to buying the full version (which yields a net utility of \( (1 + \beta \theta) n_f - p_f \)) or not buying the product. Note that \( n_f \) denotes the expected number of consumers buying the good: as choices are simultaneous, consumers base their purchasing decision on the expectations they form about future network sizes. The consumer who is indifferent between the two options is thus identified by \( n_f \), with
We suppose that consumers form rational expectations. Therefore, we have in equilibrium that $n_f^* = 1 - \theta_f$ (because all consumers with a higher valuation than $\theta_f$ decide to buy the full version), which means that

$$p_f = n_f^*(1 + \beta)(1 - n_f^*).$$

Because of the presence of network effects, there might exist more than one $n_f^*$ (that is, more than one quantity) that satisfies the equilibrium condition (6) for a given price. For instance, $p_f = 0$ for $n_f^* = 0$ and $n_f^* = 1$. We need thus a rule to choose between multiple solutions. We apply the Pareto criterion: if, at a certain price, there exist more than one number of consumers (“quantity”) that satisfies the equilibrium condition and one of these “quantities” Pareto-dominates (i.e., makes everyone better off than) the other quantities, consumers expect this allocation to prevail in equilibrium. In the present case, the larger $n_f^*$ (which corresponds to the lower $\theta_f$) that satisfies (6) gives a larger value to the network good, so everyone would be better off by coordinating on this solution. This is thus the value we pick when following the Pareto-criterion.

We can now express the firm’s problem as

$$\max_{n_f} p_f n_f = n_f^*(1 + \beta)(1 - n_f^*).$$

The first-order condition for profit maximization is $n_f^* (2 - 3 n_f^*) = 0$, which admits two roots: $n_f^* = 0$ and $n_f^* = 2/3$. Checking for the second-order condition, we find that the former corresponds to a minimum while the latter corresponds to a maximum. It follows that, when selling the full version, the firm’s optimal network size, price and profit are given by (with the superscript F meaning “full version only”):

$$n_f^F = \frac{2}{3}, \quad p_f^F = \frac{2(1 + \beta)}{9}, \quad \pi^F = \frac{4(1 + \beta)}{27}.$$
**Introduce the read-only version**

The introduction of the read-only version aims at achieving versioning. The idea is to segment the market into two segments: the full version is targeted towards consumers with a high valuation, and the read-only version towards consumers with a low valuation. We can therefore follow the same methodology as in Section 2.1. We identify two pivotal consumers. Let $\theta_f$ denote the consumer who is indifferent between the two versions. That is,

$$\theta_f \left(n_f^e + n_r^e\right) + \beta \theta_f n_f^e - p_f = \beta \theta_f n_r^e - p_r \iff \theta_f = \frac{p_f - p_r}{n_f^e + n_r^e}.$$  

Similarly, let $\theta_{ro}$ denote the consumer who is indifferent between the read-only version and no purchase:

$$\beta \theta_{ro} n_f^e - p_r = 0 \iff \theta_{ro} = \frac{p_r}{\beta n_f^e}.$$  

To achieve the desired segmentation, the firm must choose $p_f$ and $p_r$ so that $0 \leq \theta_{ro} < \theta_f < 1$. Suppose for now it is the case. Because consumers form rational expectations, the expectations about the respective network sizes have to be fulfilled at equilibrium. It follows that $n_f = 1 - \theta_f$ and $n_r = \theta_f - \theta_{ro}$. Plugging the values of $\theta_f$ and $\theta_{ro}$ into the latter two expressions, we can solve for the prices and derive the following two equilibrium conditions:

(8)  
$$p_f = \left(n_f + n_r\right)\left(1-n_f\right) + \beta n_f \left(1-n_f - n_r\right),$$  

(9)  
$$p_r = \beta n_f \left(1-n_f - n_r\right).$$  

As before, there can be multiple pairs of $(n_f, n_r)$ that meet the two conditions for a given pair $(p_f, p_r)$. It can be shown, however, that different equilibrium pairs can always be ordered, in the sense that a larger value of $n_f$ always corresponds to a larger value of $n_r$ (see Csorba and Hahn 2003). Since larger network sizes confer higher utility to all consumers, the Pareto criterion tells us that consumers expect the largest pair $(n_f, n_r)$ satisfying (8) and (9) to be the equilibrium.

We can now write the firm's profit-maximisation problem in the case where the read-only version is used as a versioning device:
To find the (unconstrained) profit-maximising values of \( n_f \) and \( n_r \), we proceed in three steps. First, solving for the FOC with respect to \( n_r \), we find that

\[
(10) \quad n_r = \left(\frac{1}{2\beta}\right) \left(1 + \beta - (1 + 2\beta)n_f\right).
\]

Second, we plug this value into the FOC with respect to \( n_f \), which then rewrites as

\[
\left(\frac{1}{4\beta}\right)(n_f - 1 - \beta)(3n_f - 1 - \beta) = 0.
\]

Third, we consider the two possible roots of the latter equation. Either \( n_f = 1 + \beta \) or \( n_f = (1 + \beta)/3 \). Using expression (10), we observe that for \( n_f = 1 + \beta \), \( n_r = -(1 + \beta) \), while for \( n_f = (1 + \beta)/3 \), \( n_r = (1 - \beta^2)/(3\beta) \).

Since we impose \( \theta_{n_r} < \theta_{n_f} \) (i.e., \( n_r > 0 \)), we can reject the former solution.

As for the latter solution, it satisfies the constraints providing (i) \( n_r > 0 \iff \beta < 1 \), and (ii) \( n_f + n_r = (1 + \beta)/(3\beta) \leq 1 \iff \beta \geq 1/2 \). Therefore, there are three cases to consider according to the value of \( \beta \) (i.e., the ratio between the valuations of the reading and writing functions).

1. **If the reading function is valued relatively higher than the writing function \( (\beta \geq 1) \)**, there is no interior solution to the above problem. The firm does not find it profitable to introduce the read-only version (which amounts to set \( n_r = 0 \)) and there is no versioning.

2. **If the reading function is valued relatively lower than the writing function \( (\beta < 1) \)**, versioning is profitable. Two cases of interest appear.

   a) **For \( 1/2 < \beta < 1 \), \( n_f = (1 + \beta)/3 \) and \( n_r = (1 - \beta^2)/(3\beta) \) meet the constraints. The corresponding optimal prices are both strictly positive and are given by (with the superscript \( \text{V} \) for “versioning”)

   \[
   p_f^\text{V} = \frac{(1 + \beta)(2\beta - 1)}{9}, \quad p_r^\text{V} = \frac{(1 + \beta)(2 - \beta)}{9\beta}.
   \]

   We compute the resulting profit as
\[ \pi^V = \frac{(1 + \beta)^3}{27 \beta}, \]

and we check that \( \pi^V > \pi^F \quad \forall 1/2 < \beta < 1. \)

b) For \( 0 < \beta \leq 1/2, \) the valuation for the reading function becomes relatively very small and the firm finds it optimal to cover the whole market (i.e., the constraint \( n_f + n_r \leq 1 \) is binding).

Since \( n_f + n_r = 1, \) it follows from (9) that \( p_r = 0: \text{the read-only version is introduced for free.} \) It also follows from (8) that \( p_f = 1 - n_f, \) meaning that the firm maximizes \( \pi = n_f (1 - n_f) \) by choosing \( n_f = 1/2 \) (which implies that \( n_r = 1/2 \) too). In sum, the optimum when \( 0 < \beta \leq 1/2 \) is (with the superscript 0 for “versioning with a free read-only version”)

\[ p_r^0 = 0, \quad p_f^0 = 1/2, \quad \pi^0 = 1/4. \]

Again, it is easily checked that \( \pi^0 > \pi^F \quad \forall 0 < \beta \leq 1/2. \)

To conclude, we summarize our findings.

Consider a software that combines a read and a write function. Suppose that consumers’ valuations for the two functions are proportional. As long as the reading function is valued relatively lower than the writing function, the seller finds it profitable to engage in versioning by selling a read-only version along with the full (read + write) version of the software. If the relative valuation of the reading function is sufficiently low, it is even profitable to give away the read-only version for free.

3.3 Conditioning prices on purchase history

In the e-commerce world, sellers are able to monitor consumer transactions, typically through the use of “cookies”. A cookie is a unique identifier which is sent by a Web site for storage by the consumer’s Web browser software. The cookie contains information about the current transaction and persists after the session has ended. As a result, at the next visit of the Web site by the consumer, the server can retrieve identification and match it with details of past interactions, which allows the seller to condition the price offers that he makes today on past behavior.
In other words, cookies make price discrimination on an individual basis feasible. Note that other technologies can be used toward the same objective: static IP addresses, credit card numbers, user authentication, and a variety of other mechanisms can be used to identify user history. Of course, users can take defensive measures. No one is forced to join a loyalty program, and it is possible to set one's browser to reject cookies or to erase them after a session is over.

In sum, online technologies allow e-commerce sellers to post prices, observe the purchasing behavior at these prices, and condition future prices on observed behavior. Yet, consumers are free to hide their previous behaviour (possibly at some cost) and can always pretend that they visit a web site for the first time. Therefore, as usual in versioning, sellers are bound to offer buyers some extra benefits in order to prevent them from hiding their identity.

We now extend the model of Section 2.1 to investigate this type of strategic interaction between buyers and sellers in an e-commerce environment. A single profit-maximizing seller provides a good at constant marginal cost (which, for simplicity, is set to zero). The seller can set cookies for recording consumers’ purchase history. However, consumers can delete cookies at some (inconvenience) cost denoted by \( \gamma > 0 \). Consumers can visit the seller’s online store in two consecutive periods. In period \( i = 1, 2 \), the utility from purchasing one unit of the good is described as in expression (1):

\[
U(\theta, s_i) = k + \theta s_i - p_i,
\]

(where, as assumed before, \( 0 < k < s_1 \)). We assume that \( s_2 > s_1 \) to capture the idea that the second unit of consumption is more valuable than the first. This might occur because the second visit to the online store is more efficient than the first one (e.g., it is easier to find one’s way through the Web pages), or because the seller offers enhanced services to second-time visitors (such as one-click shopping, loyalty rewards, targeted recommendations, ...).

Let us first examine the benchmark case where the seller sets a flat price each period, denoted by \( p \). In period \( i \), the consumer who is indifferent between buying the good and not buying is identified by \( \theta_i(p) = (p - k) / s_i \). The seller chooses thus \( p \) to maximise

\[
\pi_F = p(1 - \theta_i(p)) + p(1 - \theta_2(p)) = \frac{p}{s_1s_2} (2s_1s_2 + k(s_1 + s_2) - p(s_1 + s_2)).
\]

The optimal price and profit are easily found as

---

\( ^{16} \) This model is adapted from Acquisti and Varian (2001).
Naturally, the seller will try to take advantage of cookies and condition prices on purchase history. To do so, the seller designs the following pricing scheme: (i) the price \( p_0 \) is charged to consumers having no cookie indicating a prior visit; (ii) the price \( p_b \) is charged to consumers with a cookie indicating that they bought on a prior visit; (iii) the price \( p_n \) is charged to consumers with a cookie indicating that they did not buy on a prior visit. We investigate how to implement a solution where consumers with a high \( \theta \) consume only in the first period. We identify two pivotal consumers:

- consumer \( \theta_0 \) is indifferent between buying in the first period only and not buying at all: \( k + \theta_0 s_1 - p_0 = 0 \Leftrightarrow \theta_0 (p_0) = (p_0 - k) / s_1 \);
- consumer \( \theta_b \) is indifferent between buying in both periods (and keeping her cookie) and buying in period 1 only: \( 2k + \theta_b (s_1 + s_2) - p_0 - p_b = k + \theta_b s_1 - p_0 \Leftrightarrow \theta_b (p_b) = (p_b - k) / s_2 \).

If the proposed pricing scheme is correctly designed, the seller’s profit writes as

\[
\pi_c = p_0 (\theta_b (p_b) - \theta_0 (p_0)) + (p_0 + p_b) (1 - \theta_b (p_b))
\]

Solving for the first-order conditions, one computes the optimal prices as

\[
p_0^* = \frac{k + s_1}{2} \quad \text{and} \quad p_b^* = \frac{k + s_2}{2}.
\]

Our assumption of enhanced services at the second visit \( (s_2 > s_1) \) guarantees that \( \theta_b (p_b^*) = (s_2 - k) / 2s_2 > \theta_0 (p_0^*) = (s_1 - k) / 2s_1 \). We still need to check, however, that all consumers behave rationally.

- Consider first the consumers with \( \theta \geq \theta_b (p_b^*) \) who are supposed to prefer buying in both periods and not deleting their cookie. From the definition of \( \theta_b \) and \( \theta_0 \), we already know that these consumers are worse off if they buy in period 1 only, or if they do not buy at all. There remain two options. First, they could buy in both periods but delete their cookie, securing a net utility of \( 2k + \theta (s_1 + s_2) - 2p_0 - \gamma \) (that is, by hiding
their prior purchase, they continue to pay the low price \( p_0 \) but have to incur the cost \( \gamma \). This option is dominated if and only if

\[
p^*_b \leq p^*_0 + \gamma \iff \gamma \geq \frac{1}{2} (s_2 - s_1).
\]

Second, consumers could buy in the second period only and obtain a net utility of \( k + \theta s_2 - p_n \). This option is dominated if and only if

\[
\theta s_1 \geq p^*_0 + p^*_b - p_n - k, \text{ or } \theta \geq \frac{s_1 + s_2}{2s_1} - \frac{p_n}{s_1}.
\]

The latter condition is always satisfied for these consumers if

\[
\theta_b(p^*_n) = \frac{s_2 - k}{2s_2} \geq \frac{s_1 + s_2}{2s_1} - \frac{p_n}{s_1} \iff p_n \geq \frac{s^*_2 + s_1k}{2s_2} = p^*_0 + \frac{(s_2 - k)(s_2 - s_1)}{2s_2}.
\]

Consider next the consumers situated between \( \theta_0(p^*_0) \) and \( \theta_b(p^*_n) \) who are supposed to buy in period 1 only. Again, the definition of the two pivotal consumers guarantees that these consumers are worse off if they buy in both periods (and keep their cookie) or if they do not buy at all. We still need to check whether they are also worse off in the same two remaining options as above. First, consuming in both periods and deleting the cookie leaves them worse off if and only if

\[
k + \theta s_1 - p_0 \geq 2k + \theta (s_1 + s_2) - 2p^*_0 - \gamma, \text{ or } \theta \leq \frac{p^*_0 - k + \gamma}{s_2} = \frac{s_1 - k + 2\gamma}{2s_2}.
\]

This inequality is always satisfied for these consumers if

\[
\theta_b(p^*_n) = \frac{s_2 - k}{2s_2} \leq \frac{s_1 - k + 2\gamma}{2s_2} \iff \gamma \geq \frac{1}{2} (s_2 - s_1),
\]

which is equivalent to Condition (11). Second, consuming in period 2 only is dominated if and only if \( k + \theta s_1 - p^*_0 \geq k + \theta s_2 - p_n \), or

\[
\theta \leq \frac{p_n - p^*_0}{s_2 - s_1}.
\]
Again, all consumers in the range satisfy this condition if

\[ \theta_0(p^*_n) = \frac{s_2 - k}{2s_2} \leq \frac{p_n - p_0^*}{s_2 - s_1} \iff p_n \geq p_0^* + \frac{(s_2 - k)(s_2 - s_1)}{2s_2}, \]

which is equivalent to Condition (12).

So, as long as the cost of deleting cookies is large enough (i.e., as long as Condition (11) is met), conditioning prices on purchase history is feasible (and thus profitable). The seller sets a low price for (genuine or pretended) first visitors, \( p_0^* = (k + s_1)/2 < p_F \), and a higher price for identified second-time buyers, \( p_b^* = (k + s_2)/2 > p_F \); the difference between this price and the first-visitor price is, however, inferior to the cost of deleting cookies. Finally, the price for consumers who visit the Web site for a second time but who have not purchased earlier, \( p_n^* \), is set high enough (i.e., in accordance with Condition (12) to discourage such behavior. As the seller keeps the possibility of setting \( p_0^* = p_b^* \), fixing different prices must be profit-enhancing. We check indeed that

\[ \pi_C = \frac{k^2 + s_1s_2(s_1 + s_2) + 4ks_1s_2}{s_1s_2} = \pi_F + \frac{1}{4} \left( \frac{(s_2 - s_1)^2}{\theta^2} \right). \]

Note that conditioning prices on past purchase behavior would not be feasible if the second unit of consumption (or the second visit to the Web site) did not provide consumers with a higher value than the first unit. Indeed, if \( s_2 = s_1 \), then \( \theta_0(p_0^*) = \theta_0(p_b^*) \) and the market cannot be segmented.\(^{17}\)

We can summarize our findings as follows.

As long as the cost of deleting cookies is large enough, conditioning prices on purchase history is feasible (and thus profitable). The seller sets a low price for (genuine or pretended) first visitors, and a higher price for identified second-time buyers; the difference between this price and the first-visitor price is inferior to the cost of deleting cookies. The price for consumers who visit the Web site for a second time but who have not purchased earlier is set high enough to discourage such behavior.

\(^{17}\) The same conclusion holds when \( k = 0 \), which is reminiscent of what we observed in the general model of Section 2.2.
4 Conclusion

Price discrimination consists in selling the same product (or different versions of it) to different buyers at different prices. When sellers cannot relate a buyer’s willingness to pay to some observable characteristics, price discrimination can be achieved by targeting a specific package (i.e., a selling contract that includes various clauses in addition to price) for each class of buyers. The seller faces then the problem of designing the menu of packages in such a way that each consumer indeed chooses the package targeted for her. This practice, known as versioning (or as second-degree price discrimination), is widespread in the information economy: it is not only particularly well-suited for information goods (for which consumers’ valuations might differ widely), but it is also facilitated by the use of information technologies (which allow to create different versions of the same good at very low cost and along many possible dimensions).

In this paper, we have used a simple unified framework to expose the general theory behind versioning, and to consider a number of specific applications. In the general exposition, we have studied how to implement versioning and when it is optimal to do so. Applying the general analysis to information goods, we have shown that when the consumers’ utility for an information good can be separated along two dimensions (a “key dimension” for which consumers have different valuations, and a ‘secondary dimension’ for which all consumers have the same valuation), versioning the information good along the key dimension is the most profitable option for the monopolist.

We have then extended our theoretical framework to shed light on three specific versioning strategies used in the information economy: bundling, functional degradation and conditioning prices on purchase history. Bundling consists in selling different products as a combination package (like a word processor and a spreadsheet sold in an “office suite”). It is more profitable than separate sales if the correlation between the distributions of consumer utilities for the various goods comprised in the bundle is sufficiently negative. Under this condition, bundling induces a sufficient number of consumers to acquire a good they would not have purchased otherwise. Functional degradation is a practice by which a software firm removes some functions of its original product and sell the degraded version at a lower or zero price. For instance, Adobe sells Acrobat Reader along with Adobe Acrobat. The former software, available free of charge, is designed to view and print contents written in pdf format but is not capable of producing the contents in this format. To be able to create and edit contents, users need to purchase the latter software, which is sold at a positive price. In a model involving two-sided network effects, we have shown the following. As long as the reading function is valued (by all
users) relatively lower than the writing function, the seller finds it profitable to
engage in versioning by selling a read-only version along with the full (read +
write) version of the software. If the relative valuation of the reading function
is sufficiently low, it is even profitable to give away the read-only version for
free. Finally, we have considered the practice of conditioning prices on pur-
chase history. Indeed, online technologies allow e-commerce sellers to post
prices, observe the purchasing behavior at these prices, and condition future
prices on observed behavior. Yet, consumers are free to hide their previous
behavior (possibly at some cost) and can always pretend that they visit a web
site for the first time. Not surprisingly, the analysis revealed that the profitable
use of such practice is conditional on the cost of hiding previous purchasing
behaviour being sufficiently large.

References

tory”, mimeo, University of California, Berkeley, forthcoming in Marketing
Science.

Adams, W.J. and J.L. Yellen (1976), “Commodity bundling and the burden of

Bhargava, H.K. and V. Choudhary (2001); “Second-degree price discrimina-
tion for information goods under nonlinear utility functions”, paper pre-
sented on the Proceedings of the 34th Hawaii International Conference on
System Sciences.

innovators”, mimeo, Michigan State University.

effects: An application to software markets”, mimeo, Keele University.


economic theory: An early analysis”, Quarterly Journal of Economics 84,
268–278.

The great central banana and the small Nordic potato”, Information Eco-
nomics and Policy 14, 145–165.


