Versioning 2.0: A Product Line and Pricing Model for Information Goods under Usage Constraints and with R&D Costs

Ramnath K. Chellappa
Goizueta Business School, Emory University
Atlanta, GA 30322
ram@bus.emory.edu

Amit Mehra
Indian School of Business
Hyderabad 500032, INDIA
Amit_Mehra@isb.edu

Abstract

We model a monopolist's product line and pricing decisions wherein we relax two assumptions that are critical to understanding optimal versioning strategies for digital goods such as desktop software and mobile apps and services that impact privacy. First, through a non-monotonic utility function, we allow for the fact that consumers may not enjoy free disposal in features. Second, we endogenize the firm's initial production decision, wherein extant research assumes the highest quality of the good to be given exogenously. We observe that, even in the full information case, some highest type consumers in the market will be denied their first best quality as long as there is a finite development cost of quality. While the market is always covered in the earlier case, under information asymmetry, the monopolist may not serve the complete market *even* with zero versioning and marginal costs. An uncommon result is the evidence for quality distortion wherein the highest type gets a lower quality under information asymmetry than in the full information case. We show that a vendor's marginal cost of production and consumers' usage cost are duals; increase in either will lead to increase in the number of versions in the market. Initial development costs primarily affect only the highest types in the market by capping the highest quality produced; thus while it indirectly affects the number of versions in the market it has no bearing the variable portion of the price-quality menu. We show that extant versioning results are special cases of our model and are able to isolate the impacts of versioning costs, marginal costs and initial development costs on the optimality of versioning itself.

Key words: Versioning, information goods, mechanism design, pricing, no free disposal

1. Introduction

Early research in economics has studied feature-differentiated product line and pricing decisions of physical goods vendors (often called vertical segmentation or quality segmentation models), (Mussa and Rosen, 1978). In these models, for the same price consumers strictly prefer a higher quality good to a lower quality one and usually firms find creating a quality for each consumer type optimal except under certain conditions. Generally in such models of physical goods, the firm suffers a marginal cost as well as some quality dependent cost. The decision of an automobile company to offer a luxury and economy model of a car at the same time with differing price points can be explained through such models. In the last decade, quality differentiation for information goods has received significant attention where such strategies are called versioning strategies (Varian, 1997).

Research in Information Systems has recognized versioning to be amongst the most important strategies for information goods vendors. Such goods range from database services, where quality differentiation can be created through differential delay in access (Jain and Kannan), to music and movies where products differ in the number of bundled features (Sundararajan, 2004). The decision to version (as opposed to providing a single good) can depend on a number factors including the heterogeneity of consumers in the market and the different types of production costs (Bhargava and Choudhary, 2008). Generally however, the common theme across research on versioning is the zero marginal cost of production even if there are some differences with regard to the "costless" nature of degradation or versioning costs.

1.1. No free disposal

Irrespective of context, two critical assumptions are embedded in all extant models of versioning. First, it is generally assumed that consumers enjoy "free disposal," i.e., more of a good cannot make a consumer worse off (Mas-Colell et al., 1995). We can see that utility functions in extant work on versioning embody this non-satiation property as they are always monotonic (often linear or concave) in quality or features. However for many information goods and services this assumption does not capture reality – many of Microsoft's operating systems and software are sometimes referred to as *bloatware*, meaning that there are features in the software that are simply too much or more is not necessarily better. This is because software consumption is intrinsically associated with memory usage and hence at some point the diminishing returns from features is overtaken by the increasing cost of using it. This can be a particularly severe problem for mobile operating systems where handsets have limited capacity and memory. While Microsoft's Windows Mobile OS has always suffered from this criticism, more recently it has been reported that bloatware has crept into Google's Android OS as well (Bhargava and Choudhary, 2008, Ganapati, Sundararajan, 2004, Varian, 1997).

Such intrinsic disutility is also associated with consumption of some information services as well. Recent research points out how utility from personalization services are non-monotonic concave in services due to the in-built disutility from privacy costs (Chellappa and Shivendu, 2010). Personalization services are infeasible without sharing of personal/preference information which gives rise to privacy concerns. Hence consumers are known to only prefer a subset of the services offered even if they may be free. Indeed the assumption of free disposal is increasingly being questioned in the case of other information goods and services, "*Unlike physical goods for which "free disposal" is always an option and more is, in general, always better, service delivery is intrinsically participatory. Participation requires time commitment and physical effort on the part of consumers. Thus, there is no free disposal for service …,*" Essegaier, et al. (2002, p. 151). A prior work on software bundling also recognizes this possibility such as when some consumers may find no value for add-ins and possibly even incur a penalty cost (Dewan and

Freimer, 2003). However, there is little or scant research on mechanism design for goods with no-free-disposal (NFD) in both economics and IS research.

1.2. Initial development of quality

The production of information goods and services is accompanied by zero costs of serving an additional consumer (marginal costs) and negligible costs of degradation (versioning costs). Since the cost of copying software or other digital goods are virtually zero and as degradation mostly just involves the disabling (or non-inclusion of) a subset of functions or features, extant research has generally examined versioning against these production advantages (Chen and Seshadri, 2007). While indeed these are faithful abstractions of the real-world, none of these would be possible without the first creation of the full feature-set from which the degraded products are created and made-available. Extant research has generally ignored the impact of these initial development costs on versioning; either it assumes that infinite features can be developed costlessly or has explicitly stated that "*fixed costs of developing the highest quality are sunk, and the highest available quality is exogenously specified*" (Bhargava and Choudhary, 2008).

The goal of our research is add to the understanding of a firm's versioning strategies while accounting for these two omissions. As a result in §2 we introduce a general model from which not only can we examine the impact of NFD and initial costs but also explain extant results. In this section we first examine the full information case since the results are not obvious and as they provide for a later comparison. In §3 we analyze mechanism under information asymmetry where the firm develops a menu for self-selection. We conclude with theoretical and managerial observations in §4.

2. Model

Our model consists of a *principal* – a digital goods firm with a unique production cost structure and *agents* – consumers who face resource constraints in consuming these goods. Let $x : x \in \mathbb{R}^+$ be the number of features of the information good such that higher x implies a good of larger quality (greater number of features). The firm may costlessly damage its product of quality $\overline{x}$ to any lower quality $x \in [0, \overline{x}]$. Along the lines of Musa and Rosen (1978) and others in the versioning literature (Bhargava and Choudhary, 2008, Sundararajan, 2004, Varian, 1997), consumers are indexed with their marginal value for quality $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$ which is distributed with density function $f(\theta)$ and cumulative density $F(\theta)$ that is continuously differentiable. Further, $f(\theta)$ is assumed to be single-peaked (uni-modal) and is everywhere positive on its support such that its hazard function $h(\theta) = \dfrac{f(\theta)}{F(\theta)}$, satisfies the monotone hazard rate property. Most common distributions satisfy these standard assumptions. In order to consume the product, the customers also have must incur a resource cost. We consider a market where consumers are homogeneous in their resource-cost coefficient given by a parameter $\lambda(\lambda > 0)$. The utility for consuming a product with quality x priced at $p : p \in \mathbb{R}^+$ for a customer with index θ is:

$$U(\theta, x, p) = \theta x - \lambda x^2 - p \tag{1}$$

Note that the firm has to decide on the highest quality it must produce along with any versioning and pricing decisions. In order to endogenize this decision, we incorporate a fixed, quality-dependent cost of creating the highest quality. We assume this cost to be convex in quality and given by cx^2. Note that there is no marginal cost suffered by firm in serving additional consumers and this fixed cost is a one-time investment in creating the total number of features. For brevity, we henceforth refer to $U(\theta, x, p)$ as $U(\theta)$ only.

Please note that one intention of our work is also to reconcile the different results in versioning literature where some have suggested no versioning (or offering a single quality good to the entire market) while others recommend versioning to be always optimal. By considering a general model where consumers have usage-related costs (captured through parameter λ) and where the vendor has initial development costs (captured through parameter c) and by subsequently manipulating these parameters we can consider all of the following cases and analyze corresponding vendor strategies:

(a) Consumers enjoy **free disposal** (standard monotonic utility function) and vendor **has no** initial development cost (alternatively highest quality is exogenously specified).

(b) Consumers enjoy **free disposal** but vendor **has an** initial development cost.

(c) Consumers suffer from **no-free-disposal** but vendor **has no** initial development cost.

(d) Consumers suffer from **no-free-disposal** and vendor **has an** initial development cost.

Clearly case (d) is the most general case and we derive vendor strategies for this case throughout as this will also allow us to easily examine the other cases.

The most general understanding of versioning strategies is provided by mechanism design under information asymmetry. While this is a general starting point for most literature on versioning, we insist on providing a brief discourse on the full information case as the initial development costs are likely to create some differences. In fact as we shall see below, there is a potential for full information strategies to be different from extant models of first-degree discrimination because of the endogenization of the maximum quality decision.

2.1. Versioning Strategies under Full Information

The general timeline for the vendor in our model is that he develops a product of a certain quality level and then sells this quality and/or other reduced quality version(s) to the customers in the market. It is costless for the vendor to create version(s) of reduced quality (zero versioning costs) and he incurs no additional costs in serving the same quality to another consumer (zero marginal cost). To determine this highest quality level, the vendor has to backward induct considering his next stage decision of versions and corresponding prices.

Let $\hat{x}$ be the highest quality level that is produced by the vendor and $x(\theta)$ be the quality offered to each consumer of type θ. In the full information case, the vendor knows each consumer's type and hence he will extract the maximum surplus possible from each type. Note that our utility function is non-monotonic concave, i.e., each consumer has a satiation point at which he derives maximum benefit from consumption. This is maximized at $x^*(\theta) = \arg\max_{x(\theta)}\left[\theta x(\theta) - \lambda x^2(\theta)\right] = \frac{\theta}{2\lambda}$, and the corresponding price to extract the full surplus is given by $\mathrm{U}(\theta) = \theta x(\theta) - \lambda x^2(\theta) - p(\theta) = 0$ implying $p^*(\theta) = \frac{\theta^2}{4\lambda}$. It is very simple to observe that even if the vendor had no cost of quality, there is no point in creating a quality greater than $\frac{\overline{\theta}}{2\lambda}$ as this is the quality at which the highest type in the market derives maximum benefit from consumption. However, when there is a cost associated with quality production, we do not know if the vendor may even be able to supply this quality to the market.

Suppose if the vendor can create only the utility maximizing quality of a customer of type $\hat{\theta} \in [\underline{\theta}, \overline{\theta}]$ for whom $x^*(\hat{\theta}) = \hat{x}$. Since $x^{*\prime}(\theta) > 0$, this will imply that customer types $\theta \in (\hat{\theta}, \overline{\theta}]$ will be served quality $\hat{x}$ that is less than their first best (i.e. utility maximizing) quality. The corresponding price to extract full surplus from these customers is $p^*(\theta) = \theta\hat{x} - \lambda\hat{x}^2$. This situation is depicted in Figure 1. The corresponding objective function of the vendor is:

$$\max_{\hat{x}} \int_{\underline{\theta}}^{\hat{\theta}(\hat{x})} \frac{\theta^2}{4\lambda} f(\theta)\, d\theta + \int_{\hat{\theta}(\hat{x})}^{\overline{\theta}} \left[\theta\hat{x} - \lambda\, \hat{x}^2\right] f(\theta)\, d\theta - c\hat{x}^2 \tag{2}$$

Solving the maximization problem in (2) by Fubini's theorem and point-wise maximization we have the following Lemma.

Lemma 1. *The solution to* $\hat{\theta}$, $\hat{\theta}^*$, *is obtained by solving* $\overline{\theta} - \hat{\theta} = G(\overline{\theta}) - G(\hat{\theta}) + \frac{c\,\hat{\theta}}{\lambda}$, *where* $G(\theta) = \int F(\theta)\, d\theta$. ■

Proposition 1. *The market is covered such that the vendor provides*

(a) $x(\theta) = \frac{\theta}{2\lambda} \quad \forall \theta \in \left[\underline{\theta}, \hat{\theta}^*\right]$

(b) $x(\theta) = \hat{x} = \frac{\hat{\theta}^*}{2\lambda} \quad \forall \theta \in \left(\hat{\theta}^*, \overline{\theta}\right]$ ■

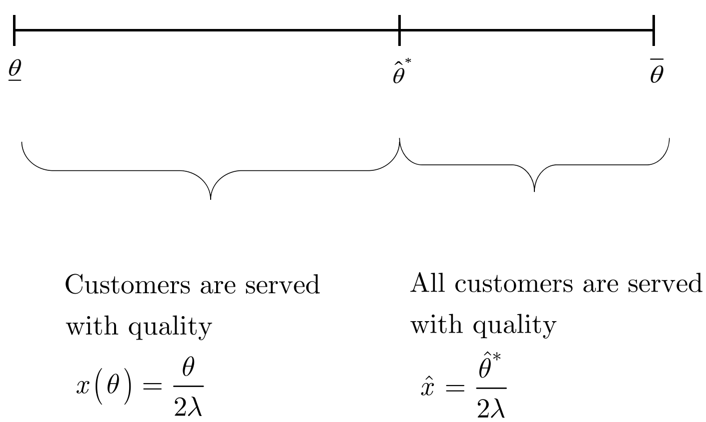


Figure 1: Market coverage and segmentation under full information

The properties of the above results make for interesting analyses. If $c = 0$ the solution to the above equation is $\hat{\theta}^* = \overline{\theta}$, i.e., all consumers will get the utility maximizing individualized version $\left(x^*(\theta)\right)$ and the highest quality that will be produced is $\frac{\overline{\theta}}{2\lambda}$. This is clearly the result for the full information analyses for case (c) and is consistent with extant full information results. However if $c > 0$ then $\hat{\theta}^* < \overline{\theta}$ since $G(\theta)$ is an increasing superlinear function of θ, i.e., consumers with index greater than $\hat{\theta}^*$ are served with quality $\hat{x}$, which is less than their utility maximizing quality. The number of versions served is reduced in this situation (case (d)) as compared to the situation defined by case (c). In other words as long as the vendor has some finite cost of creating the initial quality *even* under full information he will not offer the first best to the highest type in the market.

COROLLARY 1. *Under full information, consumers' usage-related cost is a dual of a vendor's marginal cost and is the sole reason for versioning.* ■

Note that if both $\lambda = 0$ and $c = 0$ (case (a)), we do not get interior solutions. And indeed even when $\lambda = 0$ and for some positive value of c, the solution will dictate creating the highest possible quality and serving the same quality to all types but charging different prices. Thus all extant research that has assumed a linear utility and an exogenously specified highest quality cannot have a versioning result. If any are derived, they are likely from the rent transfer engendered by information asymmetry.

In §3 we shall examine the impact of information asymmetry when such consumption and production costs are involved.

3. Versioning Strategies under Information Asymmetry

When the vendor cannot perfectly price discriminate between consumer types it must develop a menu of truth-revealing versions and prices such that the consumers self select the version targeted at them. In this case the vendor only knows of the distribution of the types and not the types themselves. Similar to the full information case, the vendor has to decide the highest quality that he will produce and the subsequent versions that he will create for the market. In determining his prices he may have to pay information rent to high types so that they are not tempted by the low quality version. We can consider these through defining the objective function of the vendor along with the respective individual rationality (IR) and incentive compatibility (IC) constraints.

Suppose if the vendor creates a highest quality x_H , the corresponding profit maximization problem for the firm is:

$$\begin{aligned} & \max_{\{x(\theta),p(\theta)\}} \int_{\underline{\theta}}^{\bar{\theta}} p\left(\theta\right) f(\theta) d\theta \\ \text{s.t.} \quad & U\left(\theta\right) \geq 0 \qquad \text{(IR)} \\ & U(\theta) \geq U_{\tilde{\theta}}(\theta) \qquad \text{(IC)} \end{aligned} \tag{3}$$

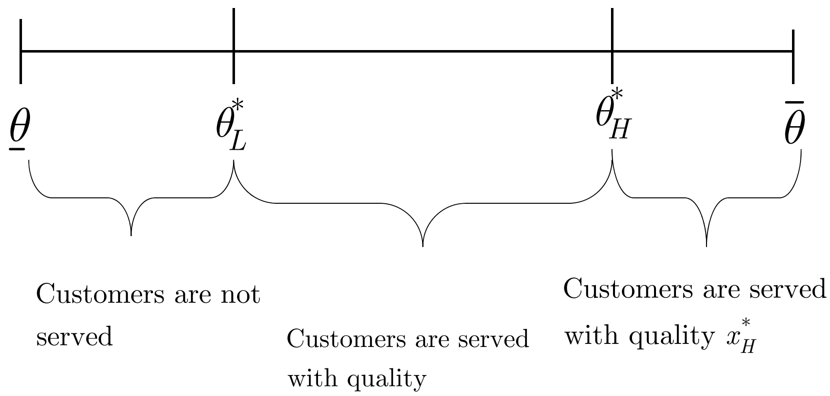
where $U_{\tilde{\theta}}(\theta)$ represents the utility of the customer of type θ if she misrepresents her type as $\tilde{\theta}$. The incentive compatibility condition essentially states that if a consumer has to pick up the price-quality meant for him then the utility from that pair should be higher than provided by any other pair meant for any other type. Please see appendix for the derivation of the truth revealing menus.

Lemma 2. *The index of the lowest customer type who is served,* θ_L^* *, is a solution to* $\theta - \left[\frac{1-F(\theta)}{f(\theta)}\right] = 0$ *and the index of the lowest customer type who gets served the highest quality,* θ_H^* *, is obtained by solving* $\theta - \left[\frac{1-F(\theta)}{f(\theta)}\right] - 2\,\lambda\, x_H = 0$. ■

Proposition 2. *The vendor serves the market such that*

$$x^*\left(\theta\right) = \begin{cases} 0 & \text{for } \theta \in \left[\underline{\theta}, \theta_L^*\right) \\ \dfrac{\theta - \left[\dfrac{1-F(\theta)}{f(\theta)}\right]}{2\lambda} & \text{for } \theta \in \left[\theta_L^*, \theta_H^*\right) \\ x_H\left(\theta_H^*\right) & \text{for } \theta \in \left[\theta_H^*, \bar{\theta}\right] \end{cases} \quad ■$$

The results provided in Lemma 2 and Proposition 2 are captured in Figure 2. This is an intermediate result in that we do not yet know what the optimal highest quality $\left(x_H^*\right)$ produced should be. Note that the monopolist develops the market into three distinct segments.



$$x\left(\theta\right) = \frac{\theta - \left[\dfrac{1-F(\theta)}{f(\theta)}\right]}{2\lambda}$$

Figure 2: Market coverage and segmentation under information asymmetry

First, he does not serve a portion of the market given by type $\theta \in \left[\underline{\theta}, \theta_L^*\right)$. While this is consistent with extant segmentation models for physical goods (with marginal costs of production) where some low types get left out of the market, this result should be somewhat surprising for information goods. Note that our monopolist suffers neither a versioning cost nor any additional cost of serving the low types in the market. In other words he could have costlessly served this segment and potentially extracted a surplus equal to $\int_{\underline{\theta}}^{\theta_L^*} p\left(\theta\right) f\left(\theta\right) d\theta$ and yet he finds it optimal not to. The economic rationale behind this decision stems from information rent that he has to pay to higher types whenever a product of lower quality-price is offered. This rent, derived from the incentive compatibility constraint, has to be paid so as to deter any temptation on the part of the high-types. The monopolist considers the tradeoff between the revenue (as there are no costs) from these low types and the net rent he has to pay to high types due to the existence of these versions and decides not to serve a segment at all.

Second, he provides a non-linear menu for a segment given $\theta \in \left[\theta_L^*, \theta_H^*\right)$ where each customer gets a version corresponding to his type $x^*\left(\theta\right)$. In other words, Proposition 2 tells us that as long as there is some positive usage-related costs to the consumer (NFD property of the good is present), the vendor will find it optimal to version. But this menu is decreasing in the consumers' usage-related cost implying that the vendor prefers to lower the quality to each type with increasing λ. For the consumer segment defined by $\theta \in \left[\theta_H^*, \overline{\theta}\right]$, the firm offers a single product. In extant segmentation models the lowest type $\left(\underline{\theta}\right)$under asymmetry is either not served at all or receives a lower quality than in the full information case. However, the highest type $\left(\overline{\theta}\right)$should generally get the same quality as in the full information case. Therefore now to solve for the complete schedule we solve for the maximum quality level the firm will produce. The objective function taking into account the pricing and versioning decisions is given as

$$\max_{x_H} \quad \int_{\theta_L^*(x_H)}^{\theta_H^*(x_H)} \left[\theta - \lambda x\left(\theta\right) - \left[\frac{1-F(\theta)}{f(\theta)}\right]\right] x(\theta) f\left(\theta\right) d\theta + \int_{\theta_H^*(x_H)}^{\overline{\theta}} \left[\theta - \lambda x_H - \left[\frac{1-F(\theta)}{f(\theta)}\right]\right] x_H f(\theta)\, d\theta - c\, x_H^2 \tag{4}$$

where $x(\theta) = \dfrac{\theta - \left[\dfrac{1-F(\theta)}{f(\theta)}\right]}{2\lambda}$. From Lemma 2, we see that θ_H^* is expressed as a function of the highest quality x_H. Hence, to get a complete clarity on θ_H^*, we would have to solve the firm's optimization problem in (4) to obtain the equilibrium highest quality x_H^*. This result in presented in Lemma 3.

LEMMA 3. *The highest quality,* x_H^*, *produced by the vendor under asymmetric information is obtained by simultaneously solving* $\theta_H^* - \left[\frac{1-F\left(\theta_H^*\right)}{f\left(\theta_H^*\right)}\right] - 2\lambda x_H^* = 0$ *and* $\theta_H^*\left[1-F\left(\theta_H^*\right)\right] = 2x_H^*\left[\lambda\left[1-F\left(\theta_H^*\right)\right]+c\right]$ ■

We are further interested in comparing the highest quality produced under full and asymmetric information and also in investigating whether customers are served with more or less than their efficient quality under the asymmetric information case. The result of this investigation is presented in Proposition 3.

PROPOSITION 3. *The highest quality under information asymmetry is lower than the highest quality produced under full information (*$\hat{x}^* > x_H^*$*).* Further, t*he optimal schedule of quality under information asymmetry is reduced as compared to full information case for every customer.*

■

This is perhaps the most important result of our work. Quality distortion implies that under information asymmetry, the highest type in the market receives a quality lower than the quality she receives under full information. Note that in conventional vertical segmentation models the highest type always receives the same quality under both full and asymmetric cases even if the price she pays in the latter case is lower (due to information rent).

We also want to develop an understanding of the market size that is covered with product versions with inferior quality than the highest quality under both full information and information asymmetry. To do this we consider the comparative statics of $\hat{\theta}^*$, θ_H^* and θ_L^* with respect to c and λ. This is our next result.

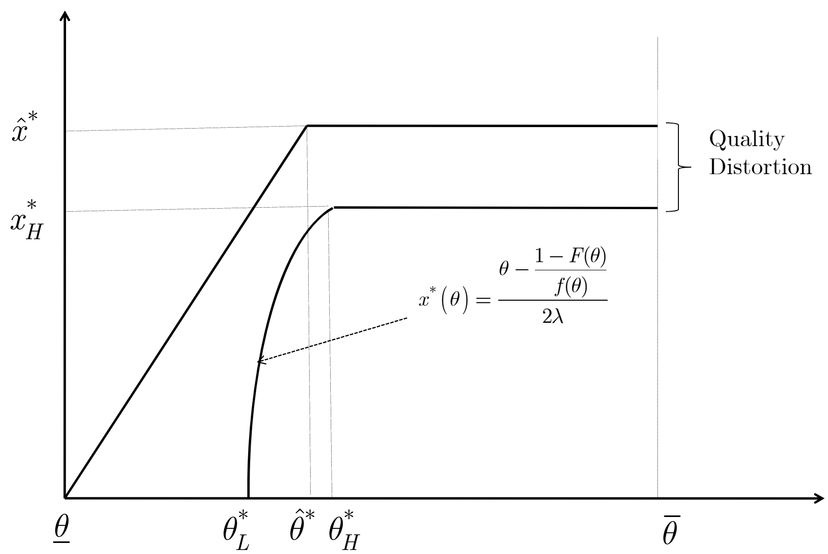


Figure 3: Quality distortion

PROPOSITION 4. *Both $\hat{\theta}^*$ and θ_H^* are decreasing in c but is increasing in λ. However, θ_L^* is independent of both these parameters.* ■

Proposition 4 succinctly captures the differential impact of production cost and the usage cost on versioning. It is increasing with λ since as the net price paid by customers reduces, the firm makes an effort to make up for the loss by producing a higher quality product. Consequently, more customers can be served with their first best quality implying that numbers of versions increases with λ. The same kind of effect is also observed under the situation of information asymmetry.

3.1. Social welfare

PROPOSITION 5. *Producer, consumer and social welfare are all decreasing in both usage and development costs.* ■

References

H. Bhargava, Choudhary, V. 2008. Research Note: When is Versioning Optimal for Information Goods? *Management Science*. **54**(5) 1029-1035.

R.K. Chellappa, Shivendu, S. 2010. Mechanism Design for "Free" but "No Free Disposal" Services: The Economics of Personalization under Privacy Concerns. *Management Science*. **56**(10) 1766-1780.

Y.-J. Chen, Seshadri, S. 2007. Product Development and Pricing Strategy for Information Goods Under Heterogenous Outside Opportunities. *Information Systems Research*. **18**(2) 150-172.

R.M. Dewan, Freimer, M.L. 2003. Consumers prefer bundled add-ins. *Journal of Management Information Systems*. **20**(2) 99-111.

S. Essegaier, Gupta, S., Zhang, Z.J. 2002. Pricing Access Services. *Marketing Science*. **21**(2) 139–159.

P. Ganapati. 2010. Bloatware Creeps Into Android Phones. *Wired.com* http://www.wired.com/gadgetlab/2010/2007/bloatware-android-phones/.

S. Jain, Kannan, P.K. 2002. Pricing of Information Products on Online Servers: Issues, Models, and Analysis. *Management Science*. **48**(9) 1123 - 1142.

A. Mas-Colell, Whinston, M.D., Green, J.R. 1995. *Microeconomic Theory*. Oxford University Press, USA.

M. Mussa, Rosen, S. 1978. Monopoly and Product Quality. *Journal of Economic Theory*. **18** 301-317.

A. Sundararajan. 2004. Nonlinear Pricing of Information Goods. *Management Science*. **50**(12) 1660-1673.

H.R. Varian. 1997. *Versioning Information Goods*. MIT Press, City.

Appendix: Proof of Lemmas & Propositions

Proof of Lemma 1

Using the analysis in Section 3.1 for the pricing and quality offered to the consumers, Expression (2) can be rewritten as:

$$\int_{\underline{\theta}}^{\hat{\theta}} \frac{\theta^2}{4\lambda} f(\theta)\, d\theta + \int_{\hat{\theta}}^{\overline{\theta}} \left[\theta \frac{\hat{\theta}}{2\lambda} - \lambda \left[\frac{\hat{\theta}}{2\lambda}\right]^2\right] f(\theta)\, d\theta - c\left[\frac{\hat{\theta}}{2\lambda}\right]^2$$

After integrating and simplifying the above expression becomes:

$$\frac{1}{4\lambda}\left[\hat{\theta}^2 F\left(\hat{\theta}\right) - \int_{\underline{\theta}}^{\hat{\theta}} 2\theta F\left(\theta\right) d\theta\right] + \frac{\hat{\theta}}{2\lambda}\left[\overline{\theta} - \hat{\theta} F\left(\hat{\theta}\right) - \int_{\hat{\theta}}^{\overline{\theta}} F\left(\theta\right) d\theta\right] - \frac{\hat{\theta}^2}{4\lambda}\left[1 - F\left(\hat{\theta}\right)\right] - \frac{c\hat{\theta}^2}{4\lambda^2}.$$

Expressing $\int F\left(\theta\right) d\theta = G\left(\theta\right)$ and $\int G\left(\theta\right) d\theta = H\left(\theta\right)$, the above expression can be further simplified to:

$$\frac{1}{4\lambda}\left[\hat{\theta}^2 F\left(\hat{\theta}\right) - 2\left[\hat{\theta} G\left(\hat{\theta}\right) - G\left(\underline{\theta}\right)\underline{\theta}\right] + 2\left[H\left(\hat{\theta}\right) - H\left(\underline{\theta}\right)\right]\right] + \frac{\hat{\theta}}{2\lambda}\left[\overline{\theta} - \hat{\theta} F\left(\hat{\theta}\right) - G\left(\overline{\theta}\right) + G\left(\hat{\theta}\right)\right]$$
$$-\frac{\hat{\theta}^2}{4\lambda}\left[1 - F\left(\hat{\theta}\right)\right] - \frac{c\hat{\theta}^2}{4\lambda^2}.$$

We represent the above expression by the symbol E and consequently $\frac{dE}{d\hat{\theta}} = \left[\overline{\theta} - \hat{\theta}\right] - \left[G\left(\overline{\theta}\right) - G\left(\hat{\theta}\right)\right] - \frac{c\hat{\theta}}{\lambda}$. The first order condition can therefore be expressed as $\overline{\theta} - \hat{\theta} = G(\overline{\theta}) - G(\hat{\theta}) + \frac{c\,\hat{\theta}}{\lambda}$. Further, $\frac{d^2E}{d\hat{\theta}^2} = -\frac{1}{2\lambda} + \frac{F\left(\hat{\theta}\right)}{2\lambda} - \frac{c}{2\lambda^2} < 0$, since $F\left(\hat{\theta}\right) \leq 1$. Thus E is strictly concave in $\hat{\theta}$ and so an internal solution is possible. Also, note that at $c = 0$, the expression $\overline{\theta} - \hat{\theta} = G(\overline{\theta}) - G(\hat{\theta}) + \frac{c\,\hat{\theta}}{\lambda}$ reduces to $\overline{\theta} - \hat{\theta} = G(\overline{\theta}) - G(\hat{\theta})$. It is easy to see that $\hat{\theta}^* = \overline{\theta}$ is a solution to this equation. Further, no other value of θ can be a solution to $\hat{\theta}$ since that can happen only when $G\left(\theta\right)$ is a linear function of θ. This would imply $F\left(\theta\right)$=1, or that all probability is a mass at one point and there is no distribution of customers. Since this is not the case, we conclude that $\hat{\theta}^* = \overline{\theta}$ is the unique solution to $\overline{\theta} - \hat{\theta} = G(\overline{\theta}) - G(\hat{\theta})$. Also note that $\left.\frac{dE}{d\hat{\theta}}\right|_{c>0,\,\hat{\theta}=\theta} < \left.\frac{dE}{d\hat{\theta}}\right|_{c=0,\hat{\theta}=\theta} \forall\theta$. Thus it must be that $\left.\hat{\theta}^*\right|_{c>0} < \left.\hat{\theta}^*\right|_{c=0}$. Thus the solution to

$\overline{\theta} - \hat{\theta} = G(\overline{\theta}) - G(\hat{\theta}) + \frac{c\,\hat{\theta}}{\lambda}$ is unique and internal. ■

Proof of Proposition 1

From Lemma 1, $\hat{\theta}^* < \overline{\theta}$. The implication is that customers with $\theta > \hat{\theta}^*$ do not receive their most efficient quality. Further, all these customers are served with the quality $\frac{\hat{\theta}^*}{2\lambda}$. The remaining customers with $\theta \in \left[\underline{\theta}, \hat{\theta}^*\right]$ are served their most efficient quality since the firm maximizes the customers' surplus and then fully extracts that surplus to maximize its profits. ■

Proof for Lemma 2 and Proposition 2

The maximization problem for the vendor is given by:

$$\begin{aligned}
&\max_{\{x(\theta),p(\theta)\}} \int_{\underline{\theta}}^{\overline{\theta}} p\left(\theta\right) f(\theta) d\theta \\
&\text{s.t.} \quad U\left(\theta\right) \geq 0 \qquad \text{(IR)} \\
&\qquad\;\; U(\theta) \geq U_{\tilde{\theta}}(\theta) \qquad \text{(IC)}
\end{aligned}$$

We first focus on the IC condition. Suppose the vendor offers a quality/feature-price schedule $\left\{x\left(\theta\right), p\left(\theta\right)\right\}$ for every type θ. To make sure that the customers self-select into buying the appropriate version, it must be that each customer maximizes her surplus by truthfully revealing her type θ. In other words, the customers' incentive compatibility constraints (ICs) must be satisfied. We represent the utility of a customer of type θ who declares her type to be $\tilde{\theta}$ as $U_\theta(\tilde{\theta})$. Hence, it must be that:

$$U_\theta(\theta) \geq U_\theta(\tilde{\theta}) \Rightarrow \theta x(\theta) - \lambda x^2(\theta) - p(\theta) \geq \tilde{\theta} x(\tilde{\theta}) - \lambda x^2(\tilde{\theta}) - p(\tilde{\theta}) \tag{5}$$

for any $(\theta, \tilde{\theta}) \in [\underline{\theta}, \overline{\theta}] \times [\underline{\theta}, \overline{\theta}]$. Similarly, for a customer of type $\tilde{\theta}$, it must be true that declaring herself to be of type θ would result in lower utility for her. Corresponding to Equation (5), we get

$$U_{\tilde{\theta}}(\tilde{\theta}) \geq U_{\tilde{\theta}}(\theta) \Rightarrow \tilde{\theta} x(\tilde{\theta}) - \lambda x^2(\tilde{\theta}) - p(\tilde{\theta}) \geq \tilde{\theta} x(\theta) - \lambda x^2(\theta) - p(\theta) \tag{6}$$

Adding equations (5) and (6), we get

$$[x(\theta) - x(\tilde{\theta})][\theta - \tilde{\theta}] \geq 0 \tag{7}$$

Thus the incentive-compatibility constraint requires that the schedule of features $x(\theta)$ has to be non-decreasing, i.e.,

$$x'(\theta) \geq 0 \tag{8}$$

Further, incentive compatibility also implies that truthful revelation of one's type would result in utility maximization. Thus, for a customer of type θ, it must be that $\left.\frac{dU_\theta(\tilde{\theta})}{d\tilde{\theta}}\right|_{\tilde{\theta}=\theta} = 0$ because of the appropriate first order conditions. This is simplified as:

$$\theta x'(\theta) - 2\lambda x(\theta) x'(\theta) - p'(\theta) = 0 \tag{9}$$

For Equation (9) to be meaningful, the utility function $U_\theta(\tilde{\theta})$ must also satisfy the second order condition, i.e., $\left.\frac{d^2 U_\theta(\tilde{\theta})}{d\tilde{\theta}}\right|_{\tilde{\theta}=\theta} < 0$. This requirement can be simplified to:

$$\theta x''(\theta) - 2\lambda\left[x'^2(\theta) + x(\theta) x''(\theta)\right] - p''(\theta) < 0 \tag{10}$$

Differentiating Equation (9) with respect to θ, we get

$$x'(\theta) + \theta x''(\theta) - 2\lambda\left[x'^2(\theta) + x(\theta)x''(\theta)\right] - p''(\theta) = 0 \quad (11)$$

Substituting from Equation (11) in (10) we obtain $x'(\theta) \geq 0$. From Equation (8), we know that this condition is required for truth revelation. Thus the second order conditions do not impose any further constraints. In order for local ICs to satisfy globally, we need that the crossing property or Spence-Mirrlees Condition to be satisfied. Since, the cross-derivative $\left(\frac{\partial^2 U(x,p,\theta)}{\partial x \partial \theta} = \frac{\partial(\theta - 2\lambda x)}{\partial \theta} = 1\right)$ has a constant sign, the requisite conditions are met.

Next, we simplify the objective function utilizing the conditions imposed by the Incentive Compatibility constraint and expressed in Equation (9). Note that:

$$U(\theta) = \theta x(\theta) - \lambda x^2(\theta) - p(\theta) \quad (12)$$

Differentiating both sides of the above equation with respect to θ, we get:

$$U'(\theta) = x(\theta) + \theta x'(\theta) - 2\lambda x(\theta)x'(\theta) - p'(\theta) \quad (13)$$

Utilizing Equation (9), we can simplify Equation (12) to

$$U'(\theta) = x(\theta) \quad (14)$$

Integrating Equation (14) between the limits $\underline{\theta}$ and θ, we get $U(\theta) - U(\underline{\theta}) = \int_{\underline{\theta}}^{\theta} x(y)dy$. Since the participation constraint of the lowest-type consumer must bind, we have $U(\underline{\theta}) = 0$. Hence, we have

$$U(\theta) = \int_{\underline{\theta}}^{\theta} x(y)dy \quad (15)$$

Using Equations (12) and (15), we can write $p(\theta) = \theta x(\theta) - \lambda x^2(\theta) - \int_{\underline{\theta}}^{\theta} x(y)dy$. Thus, we can now rewrite the vendor's objective function to

$$\int_{\underline{\theta}}^{\overline{\theta}} \left[\theta x(\theta) - \lambda x^2(\theta)\right] f(\theta)\, d\theta - \int_{\underline{\theta}}^{\overline{\theta}} \left[\int_{\underline{\theta}}^{\theta} x(y)\, dy\right] f(\theta)\, d\theta \tag{16}$$

Using Fubini's theorem we get $\int_{\underline{\theta}}^{\overline{\theta}} \left[\int_{\underline{\theta}}^{\theta} x(y)\, dy\right] f(\theta)\, d\theta = \left[\left[\int_{\underline{\theta}}^{\theta} x(y)\, dy\right] F(\theta)\right]_{\underline{\theta}}^{\overline{\theta}} - \int_{\underline{\theta}}^{\overline{\theta}} F(\theta) x(\theta)\, d\theta$. Using the fact that $F(\overline{\theta}) = 1$ and $F(\underline{\theta}) = 0$, we can simplify the right hand side of the above equation to $\int_{\underline{\theta}}^{\overline{\theta}} [1 - F(\theta)] x(\theta)\, d\theta$. Thus we can further simplify the expression in (16) to

$$\int_{\underline{\theta}}^{\overline{\theta}} \left[\theta - \lambda x(\theta) - \frac{1 - F(\theta)}{f(\theta)}\right] x(\theta) f(\theta)\, d\theta \tag{17}$$

At this point, we ignore the constraints and do an unconstrained optimization. We later check that the constraints are satisfied. By employing point-wise maximization we need to only maximize the integrand with respect to $x(\theta)$. This gives

$$x^*(\theta) = \frac{\theta - \left[\frac{1 - F(\theta)}{f(\theta)}\right]}{2\lambda} \tag{18}$$

We can now analyze the quality menu used to serve the market using Equation (18). Further, the quality being served increases with the customer index until the highest possible quality $\hat{x}$ is reached (since $\left[\frac{1 - F(\theta)}{f(\theta)}\right]$ is decreasing in θ. Hence $x^*(\theta)$ is increasing in θ which is exactly what we need to satisfy the constraint specified in Equation (8)).

Note that the marginal customer who is served gets a quality of 0. Let this customer be indexed by θ_L^*. Then we have:

$$\theta - \left[\frac{1 - F(\theta)}{f(\theta)}\right] = 0 \tag{19}$$

The solution to the above equation, θ_L^*, provides the index of the lowest type of customer who is served. Let the index of the lowest customer type who is served with full quality be given by θ_H^*. This point is the solution to:

$$x_H = \frac{\theta - \dfrac{1 - F(\theta)}{f(\theta)}}{2\lambda}$$

$$\text{or, } \theta - \frac{1 - F(\theta)}{f(\theta)} - 2\lambda x_H = 0 \tag{20}$$

Finally, note that $\theta_H^* > \theta_L^*$ since the terms in the equations (19) and (20) are identical except for an additional negative constant term in Equation (20). Hence versioning is optimal when customers suffer from No Free Disposal. ■

Proof of Lemma 3

The first order condition of Equation (4) with respect to x_t yields

$$\int_{\theta_H^*}^{\overline{\theta}} \left[\theta - 2\lambda x_H^* - \left[\frac{1 - F(\theta)}{f(\theta)}\right]\right] f(\theta)\, d\theta = 2c x_H^*$$

This can be simplified to:

$$\theta_H^*\left[1 - F\left(\theta_H^*\right)\right] = 2x_H^*\left[\lambda\left[1 - F\left(\theta_H^*\right)\right] + c\right] \tag{21}$$

Substituting θ_H^* in place of θ in Equation (20) and solving it simultaneously with Equation (21) we obtain θ_H^* and x_H^*.

Proof of Proposition 3

We need to prove that optimal highest quality in the full information case is greater than the optimal highest quality in the incomplete information situation. We represent the objective function of the vendor under complete information (Expression (2)) as a function of the highest quality x by $O_C\left(x\right)$. Hence, we have:

$$\left.\frac{dO_C(x)}{dx}\right|_{x=x_H^*} = \int_{\hat{\theta}}^{\bar{\theta}} \left[\theta - 2\lambda x_H^*\right] f\left(\theta\right) d\theta - 2cx_H^*$$

Substituting the value of $2cx_H^*$ from Equation (21) in the above equation, we get:

$$\left.\frac{dO_C(x)}{dx}\right|_{x=x_H^*} = \int_{\hat{\theta}}^{\bar{\theta}} \left[\theta - 2\lambda x_H^*\right] f\left(\theta\right) d\theta - \int_{\theta_H^*}^{\bar{\theta}} \left[\theta - \frac{1-F\left(\theta\right)}{f\left(\theta\right)}\right] f\left(\theta\right) d\theta + 2\lambda x_H^* \left[1 - F\left(\theta_H^*\right)\right].$$

Using the fact that $x_H^* = \dfrac{\hat{\theta}^*}{2\lambda}$, the above equation can be easily simplified to:

$$\left.\frac{dO_C(x)}{dx}\right|_{x=x_H^*} = \int_{\hat{\theta}}^{\theta_H^*} \left[\theta - \hat{\theta}\right] f\left(\theta\right) d\theta + \int_{\theta_H^*}^{\bar{\theta}} \left[1 - F\left(\theta\right)\right] d\theta \qquad (22)$$

Note that $\hat{\theta}^*(x_H^*) = 2\lambda x_H^*$ and from Lemma 2, $\theta_H^*(x_H^*) = 2\lambda x_H^* + \dfrac{1 - F\left(\theta_H^*\left(x_H^*\right)\right)}{f\left(\theta_H^*\left(x_H^*\right)\right)}$. Clearly, it must be that $\theta_H^*\left(x_H^*\right) > \hat{\theta}^*\left(x_H^*\right)$. This implies that the first term on the right hand side of Equation (22) must be positive. Also, the second term must be positive since $F\left(\theta\right) < 1$ for $\theta \leq \bar{\theta}$. Thus, we have shown that $\left.\dfrac{dO_C(x)}{dx}\right|_{x=x_H^*} > 0$. Further, we know that $\left.\dfrac{dO_C(x)}{dx}\right|_{x=\hat{x}^*} = 0$ and that $O_C(x)$ is a concave function in x. Thus, $\hat{x}^* > x_H^*$. This completes the proof.

Proof that customers suffer a quality distortion on the low side under incomplete information

(A) Consider $\theta < \text{Min}\left\{\hat{\theta}^*, \theta_H^*\right\}$

From Proposition 1, the quality served under full information is $x^*(\theta) = \frac{\theta}{2\lambda}$ and from Proposition 2, the quality served under incomplete information is $x^*(\theta) = \frac{\theta - \frac{1-F(\theta)}{f(\theta)}}{2\lambda}$. Since $F(\theta) < 1$, it is obvious that the quality served under information asymmetry is lower.

(B) Consider $\theta > \text{Max}\left\{\hat{\theta}^*, \theta_H^*\right\}$

All such customers are served with quality $\hat{x}^*$ under full information and x_H^* under information asymmetry. We already proved that $\hat{x}^* > x_H^*$ above. Hence, again, lower quality is served under information asymmetry.

(C) Consider $\text{Min}\left\{\hat{\theta}^*, \theta_H^*\right\} \leq \theta \leq \text{Max}\left\{\hat{\theta}^*, \theta_H^*\right\}$

Suppose $\hat{\theta}^* < \theta_H^*$. So all customer in this range will be served quality $\hat{x}^*$ under full information and a quality less than x_H^* under information asymmetry since $x^*(\theta)$ is increasing (since $\frac{1-F(\theta)}{f(\theta)}$ is decreasing in θ). Further, $\hat{x}^* > x_H^*$. Hence a reduced quality is served under information asymmetry.

Suppose $\hat{\theta}^* > \theta_H^*$. The customer indexed by θ_H^* will be served quality x_H^* under information asymmetry. Because of logic similar to (A) above, this customer must be served a higher quality under full information. Further, as θ increases, the quality under information asymmetry remains at x_H^* whereas the quality served under full information increases (since $\frac{\theta}{2\lambda}$ is increasing in θ). Hence all customers in this range are served a reduced quality under information asymmetry.

Hence proved.

Proof of Proposition 4

Comparative Statics of $\hat{\theta}^*$

Differentiating the equation from Lemma 1, $\overline{\theta} - \hat{\theta}^* = G(\overline{\theta}) - G(\hat{\theta}^*) + \frac{c\,\hat{\theta}^*}{\lambda}$ with respect to c, we get $-\frac{d\hat{\theta}^*}{dc} + \frac{dG\left(\hat{\theta}^*\right)}{d\hat{\theta}^*}\frac{d\hat{\theta}^*}{dc} - \frac{\hat{\theta}^*}{\lambda} - \frac{c}{\lambda}\frac{d\hat{\theta}^*}{dc} = 0$. This can be rewritten as $\frac{d\hat{\theta}^*}{dc}\left[1 - F\left(\hat{\theta}^*\right) + \frac{c}{\lambda}\right] = -\frac{\hat{\theta}^*}{\lambda}$.

Since $F\left(\hat{\theta}^*\right) \leq 1$, it must be that $\frac{d\hat{\theta}^*}{dc} < 0$.

Similarly, differentiating the equation in Lemma 1 with respect to λ, we get $\frac{d\hat{\theta}^*}{d\lambda}\left[1 - F\left(\hat{\theta}^*\right) + \frac{c}{\lambda}\right] = \frac{c\hat{\theta}^*}{\lambda^2}$. From this, we can easily see that $\frac{d\hat{\theta}^*}{d\lambda} > 0$.

Comparative Statics of θ_L^*

From Lemma 2, it is easy to see that θ_L^* does not depend on either c or λ.

Comparative Statics of θ_H^*

θ_H^* is obtained by solving $\theta_H^* - \frac{1 - F\left(\theta_H^*\right)}{f\left(\theta_H^*\right)} - 2\lambda x_H^* = 0$ (see Lemma 3).

Clearly, as λ increases, the solution to the above equation, i.e., θ_H^* increases. Also, as c increases, x_H^* reduces and hence θ_H^* reduces.

This concludes the proof.