# 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
ramnath.chellappa@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 version 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. We categorically prove that such marginal costs are the *sole* reason to pursue a versioning strategy. 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. Finally, we show that extant versioning results are special cases of our model and are able to isolate the impacts of 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; usually firms find creating a quality for each consumer type optimal except under certain conditions (sometimes called shutdown 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 luxury and economy models at the same time with differing feature sets and price points is explained through such models. Along these lines, in the last decade, quality differentiation for *information goods* or *digital products* has received significant attention where such product line strategies are called *versioning* (Varian 1997).

Research in information systems (IS) has recognized versioning to be amongst the most important strategies for digital 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). It has been suggested that, for a monopolist, the optimality of versioning strategy (as opposed to offering a single version of the product) can be a function of a number factors including the distributional assumptions and production costs amongst others; although many papers in IS have examined the versioning problem, the different setup conditions often do not allow for an easy reconciliation amongst the results. For example, a recent work by Jones and Mendelson (2011 p. 166) suggests that the *multiproduct efficient solution for information goods is to offer a single product* basically *because the losses from cannibalization always outweigh the benefits of segmentation.* In other words, this work suggests that versioning is a sub-optimal strategy for a monopolist even if marginal costs are zero – although the authors note that this result is conditional on certain assumptions regarding the boundary of the distribution of consumer valuations. On the other hand Bhargava and Choudhary (2008) derive conditions wherein under certain cost (to the firm) to value (to the consumer) ratios, versioning might be optimal. However while quality is an endogenous decision in the former, no quality price menu is derived in the latter work even if distributional assumptions are non-restrictive and general.

Generally, a consistent and binding aspect amongst research on versioning is the assumption of negligible marginal cost of production, i.e., the cost of serving an additional consumer is taken to be zero; variants include the use of a constant quality-independent cost of production (Bhargava and Choudhary 2008). It is to be noted that extant models often vary in their assumptions regarding the "costless" nature of degradation or versioning costs, sometimes there is simply no mention of these costs. Versioning costs are distinct and different from marginal costs of production and may be subsumed in the same assumption only in the case where there is only one consumer of each type and where each type gets a distinct version. Also note that in many models where only two consumer types are considered, perhaps using a point distribution, the versioning results do not easily and fully translate to situations with continuous types of consumers. The simple reason is that when costs of versioning are considered, the decision is two-fold; first if versioning is optimal at all and second, the number of versions to offer. In sum, it would not be an exaggeration to say that a lucid understanding of the impact of various costs on versioning decisions is much desired. This desire is underscored by the fact that independent of other initial setup conditions, two critical assumptions are embedded in most if not 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, Whinston et al. 1995) and second, the creation of the highest quality from which subsequent versions are derived is exogenously assumed, i.e., the impact of initial investments or capital costs is summarily ignored. Consistent with our previously expressed goals, in this paper we seek to examine the impact of these two additional costs on versioning and pricing decisions of an information goods vendor.

### 1.1. No free disposal

Many models in economics that employ utility functions generally assume a free disposal property (Mas-Colell, Whinston et al. 1995) implying that for the same price consumers will prefer more of the good. Utility functions in extant work on versioning also embody this non-satiation property and these functions monotonic (often linear or concave and increasing) in quality or features. However for many information goods and services this assumption does not capture reality, e.g., Microsoft's operating systems and software are sometimes referred to as *bloatware*, as they are often packed with excessive features. For these products, more is not necessarily better because software consumption is intrinsically associated with memory usage and hence at some point the diminishing return from features is overtaken by the increasing cost of using it. Such excess can be a particularly severe problem for mobile operating systems where handsets and touchpads 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 (Milian 2010). Generally, this aspect of software consumption has been ignored with the exception of one work on software bundling that 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).

Recent research points out how utility from personalization services are also 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 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). However, there is little or scant research on mechanism design for goods with no-free-disposal (NFD) in both economics and IS research. Specifically we do not how these costs suffered by the consumer will impact versioning and other product line decisions.

### 1.2. Initial development of quality

Since the cost of copying software or other digital goods is virtually zero and as degradation often just involves the disabling (or non-inclusion of) a subset of functions or features, prior 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 on information goods has 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). In the information goods context, one recent work (Jones and Mendelson 2011) considers development costs and observes the phenomenon of quality distortion, although certain unique assumptions (discussed later) in the model does not consider the monopolist's versioning strategy. In this regard, there is also a work (Hahn 2000) in the physical goods quality segmentation literature that examines the impact of initial fixed costs of developing the highest quality although the firm also suffers a marginal cost of serving each customer. As discussed later. our work follows this line of reasoning for information goods and allows for the reconciliation of the impact of production cost (suffered by the firm) and usage cost (suffered by the consumer).

In summary, an important goal of this research is to provide an overall understanding of the impact of different types of costs on versioning and other product-line decisions of the information goods vendor. In particular this paper will provide additional insights into the role of three costs – consumption and initial R&D development costs, hitherto unexamined, and versioning costs which was often co-mingled with marginal costs of production.

In §2 we introduce a general model developed along the lines of a standard monopolistic screening model (Laffont and Martimort 2001) from which not only can we examine the impact of NFD and initial costs but also relate to 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 (Varian 1997, Sundararajan 2004, Bhargava and Choudhary 2008), consumers are indexed with their marginal value for quality .. 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 our model, in order to consume the product, the customers must also incur a resource cost. For example, this might be the average memory consumed to run each software feature and we consider a market that is homogeneous in its resource-cost coefficient given by a parameter $\lambda(\lambda > 0)$. Therefore the utility for consuming a product with quality $x$ priced at $p : p \in \mathbb{R}^+$ for a customer with index $\theta$ is:

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

Observe that the utility function is non-monotonic in features, i.e. up to a point (utility maximizing set of features) the utility is increasing in features and then decreases. In other words there is no free disposal in features; more features can actually make the consumer worse-off. For brevity, we henceforth refer to $U(\theta, x, p)$ as $U(\theta)$ only.

A second aspect of our model is 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 $cx^2$. This convex cost function is commonly assumed for information goods since it generally believed that most cost-effective decisions are made first and it becomes increasingly costlier to improve quality of/add features to information goods (Jones and Mendelson 2011). Empirical research in software engineering also finds this cost to be convex although there are some differences in the degree of convexity (Boehm, Abts et al. 2000). This fixed cost is a one-time investment in creating the maximum number of features, a kind of research and development (R&D) investment. Once this list of features is created, the firm can create versions of the good. At this time we assume versioning costs to be zero (aka costless damaging) although we specifically relax this assumption later to delineate the differential impact of versioning costs and marginal costs of production. Note that we assume the latter to be negligible throughout the paper, i.e., firms can serve additional consumers for free once a version is created.

One intention of our work is also to be able to reconcile the different results in versioning literature where some have suggested the single product strategy (no versioning) to be optimal. By considering a general model where consumers have usage-related costs (captured through parameter $\lambda$) and where the vendor has initial development costs (captured through parameter $c$) and by subsequently manipulating these parameters we can consider many different setups from extant models including the standard monotonic utility function and exogenously specified highest-quality.

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

### 2.1. Versioning Strategies under Full Information – Welfare Maximizing Solution

The timeline for the model is as follows: The vendor invests in research and development to create a product of certain quality level defined by the maximum number of features he produces. From this feature set he can create other reduced quality version(s) and sets price(s) accordingly. 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.

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 the solution to this problem is the same as a welfare-maximizing solution except that in the monopolist's case, the vendor extracts all the surprlus, leaving the consumer with zero welfare. 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 $\theta$. 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 $\hat{\theta} \in \left[\underline{\theta}, \overline{\theta}\right]$ for whom $x^*(\hat{\theta}) = \hat{x}$. Since $x^{*\prime}(\theta) > 0$, this will imply that customer types $\theta \in \left(\hat{\theta}, \overline{\theta}\right]$ 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 $p^*(\theta) = \theta\hat{x} - \lambda\hat{x}^2$. The corresponding objective function of the vendor is:

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

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* $\overline{\theta} - \hat{\theta} = G(\overline{\theta}) - G(\hat{\theta}) + \dfrac{c\,\hat{\theta}}{\lambda}$ , *where* $G(\theta) = \int F(\theta)\, d\theta$ . ■

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

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

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 $\left(x^*\left(\theta\right)\right)$ and the highest quality that will be produced is $\dfrac{\overline{\theta}}{2\lambda}$, if there are no costs of production but consumers incur a usage cost. However if $c > 0$ then $\hat{\theta}^* < \overline{\theta}$ since $G(\theta)$ is an increasing superlinear function of $\theta$ , 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 to the marketplace is reduced in the latter case as compared to the situation where $c = 0$ . 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.

Note that if both $\lambda = 0$ and $c = 0$ , we do not get interior solutions. This should be fairly obvious in that if the consumers' utility is strictly increasing in features and there is no cost to producing them, then an infinite quality/price would be the solution for all consumer types. 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 – clearly an unworkable monopolist strategy (ability to sell the same product at different prices to different consumers) that is only possible under extreme conditions of discrimination. It is indeed such a result, one of offering a single quality, that Jones and Mendelson (2011) allude to in their welfare-efficient solution (which is the same as the surplus extracting monopolist solution except prices are set to zero). Our setting uses a general distribution, so clearly it is not the uniform distribution of consumer types assumed in the aforementioned model that drives this result. Before we delve deeper into any sources of versioning, in §3, we shall first 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 a lower 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 some maximum quality $x_H$, the corresponding profit maximization problem for the firm is:

$$\begin{array}{lll} \max\limits_{\{x(\theta),p(\theta)\}} & \int_{\underline{\theta}}^{\overline{\theta}} p(\theta) f(\theta) d\theta - c\left[x_H\right]^2 & \\ \text{s.t.} & U(\theta) \geq 0 & \text{(IR)} \\ & U(\theta) \geq U_{\tilde{\theta}}(\theta) & \text{(IC)} \end{array} \tag{3}$$

where $U_{\tilde{\theta}}(\theta)$ represents the utility of the customer of type $\theta$ 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. Hence, it must be that:

$$U_\theta\left(\theta\right) \geq U_\theta\left(\tilde{\theta}\right) \Rightarrow \theta x\left(\theta\right) - \lambda x^2\left(\theta\right) - p\left(\theta\right) \geq \theta x\left(\tilde{\theta}\right) - \lambda x^2\left(\tilde{\theta}\right) - p\left(\tilde{\theta}\right) \tag{4}$$

for any $\left(\theta,\tilde{\theta}\right) \in \left[\underline{\theta},\overline{\theta}\right] \times \left[\underline{\theta},\overline{\theta}\right]$. Similarly, for a customer of type $\tilde{\theta}$, it must be true that declaring herself to be of type $\theta$ would result in lower utility for her, i.e., we need that

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

This leads to understand that any optimal versioning menu, if one exists, needs to non-decreasing in consumer types (proof in the appendix), i.e.,

$$x'\left(\theta\right) \geq 0 \tag{6}$$

This gives us Lemma 2.

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

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

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

This is an intermediate result in that we do not yet know what the optimal highest quality $\left(x_H^*\right)$ produced should be. First, note that the monopolist develops the market into three distinct segments.

And 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 can be considered 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 $\int_{\underline{\theta}}^{\theta_L^*} p(\theta) f(\theta) d\theta$ 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 develops a non-linear menu for a segment given by $\theta \in \left[\theta_L^*, \theta_H^*\right)$ where each customer gets a version corresponding to his type $x(\theta)$. In other words, Proposition 2 tells us that not only will the vendor develop three distinct consumer segments but he will find it optimal to engage in a type-dependent versioning strategy for the middle segment. We can easily see that this quality menu is decreasing in $\lambda$ meaning that with increasing usage-related costs, the each consumer type's quality is lowered. 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 $(\underline{\theta})$ under asymmetry is either not served at all or receives a lower quality than in the full information case. However, the highest type $(\overline{\theta})$ 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

$$\max_{\{x_H\}} \int_{\theta_L^*(x_H)}^{\theta_H^*(x_H)} \left[\theta - \lambda x(\theta) - \left[\frac{1-F(\theta)}{f(\theta)}\right]\right] x(\theta) f(\theta)\, d\theta + \int_{\theta_H^*(x_H)}^{\bar{\theta}} \left[\theta - \lambda x_H - \left[\frac{1-F(\theta)}{f(\theta)}\right]\right] x_H f(\theta)\, d\theta - c\, x_H^2 \qquad (7)$$

where $x(\theta) = \dfrac{\theta - \left[\dfrac{1-F(\theta)}{f(\theta)}\right]}{2\lambda}$. From Lemma 2, we see that $\theta_H^*$ is expressed as a function of the highest quality $x_H$. Hence, to get a complete clarity on $\theta_H^*$, we would have to solve the firm's optimization problem in (7) 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* $\theta_H^* - \left[\dfrac{1-F(\theta_H^*)}{f(\theta_H^*)}\right] - 2\lambda x_H^* = 0$ *and* $\theta_H^*\left[1-F(\theta_H^*)\right] = 2x_H^*\left[\lambda\left[1-F(\theta_H^*)\right] + c\right]$ ■

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

■

Lemma 3 provides us the lower bound of the high-type consumers who will receive a single version as well as the quality of this version. We can now compare this highest quality that will be developed under information asymmetry with the welfare-maximizing or full-information solution given by Proposition 1. This comparison is given in Proposition 3 which points to quality distortion for the high types. 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). An exception to these standard physical goods models is a recent unpublished thesis (Hahn 2000) that also observes such distortion for all consumer types even when there is a marginal cost of production. The single version result of Jones and Mendelson (2011) also alludes to the depression in quality and although they refer to this as a property of information goods, we shall show later that this property is a function of the initial development cost, independent of the physical/digital nature of the good.

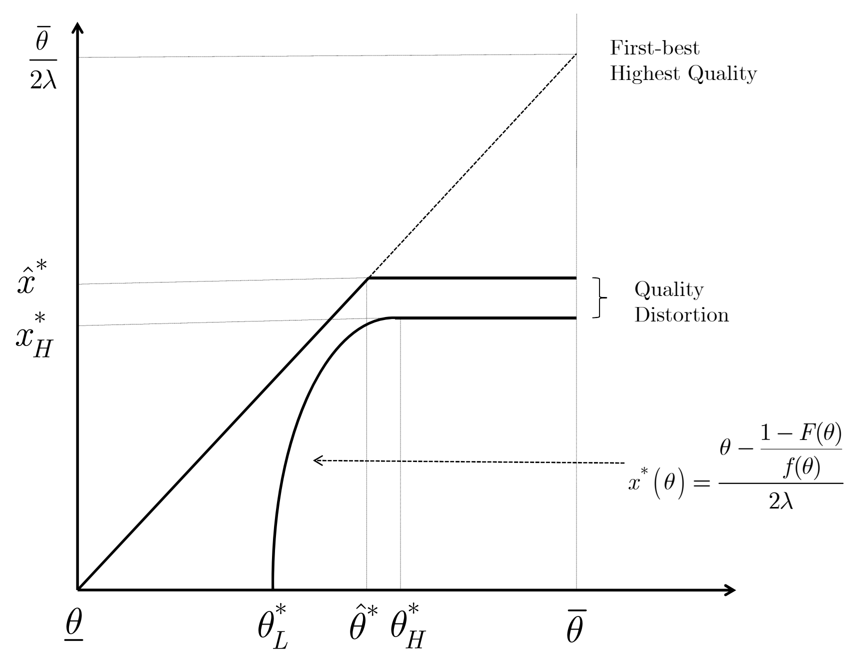


**Figure 1: Quality distortion**

Note from Figure 1 that the quality enjoyed by the highest type in the market has undergone two depressions – in other words it is the third-best quality under information asymmetry. Also note that the versioning menu is concave in consumer types while it is linear in the full information case. We know that the monopolist not only finds it optimal to divide the market into three segments but also that versioning is an optimal strategy independent of the distribution of consumer types. However, we do not know if the same can be said of market coverage. Note that in many other extant models of versioning, either explicit assumptions on market coverage are introduced during the model setup or only two types are considered; in the latter if versioning is pursued market is always covered. In our case, as a result of our mechanism-design approach to deriving the full versioning-menu for a general distribution, we endogenize any market coverage decision. We can further understand market coverage though a comparative statics of the various bounds derived in the model with respect to costs $c$ and $\lambda$.

**PROPOSITION 4.** *The number of versions are decreasing in the development costs $c$ but are increasing in the usage costs $\lambda$. Market coverage is independent of both these parameters.* ■

Proposition 4 succinctly captures the differential impact of the initial development cost and the usage cost on versioning providing key insights into market coverage and versioning as a strategy. First consider the comparative statics of $\theta_L^*$ with respect to the two costs; $\theta_L^*$ refers to the lowest consumer type who would be offered a version and from Lemma 2 we can see that this boundary is independent of either cost parameters. In other words, under information asymmetry, market coverage when versioning is optimal is purely a function of the distribution. The market is covered if an only if $\underline{\theta} = \dfrac{1}{f(\underline{\theta})}$. Clearly in the full information case, the market is always covered with the lower types receiving their first-best quality.

Now consider the comparative static of $\hat{\theta}^*$ and $\theta_H^*$; we can see that both these bounds are decreasing in $c$ but are increasing in $\lambda$. $\theta_H^*$ refers to the highest consumer type who will be served a personalized version (second-best quality) under information asymmetry while $\hat{\theta}^*$ represents such consumer type (who gets his first-best quality) under full information. Not surprisingly these bounds are decreasing in the development cost $c$ as the primary contributor to the development cost is the highest quality produced which corresponds to the personalized version of $\theta_H^*$. However it might be somewhat surprising to note that these bounds are increasing in the consumer's usage cost $\lambda$, also implying that more people get their second-best quality as usage costs increase. The economic intuition behind this observation is that usage costs take away from the surplus enjoyed by the consumer and hence the net price that the vendor can charge; the firm therefore attempts 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 $\lambda$. Indeed we see this under full information as well.

**Lemma 4.** *Versioning is optimal only if the marginal rent to the firm, from at least some consumer types, is strictly concave in features*

**Proposition 5.** *For a monopolist, marginal costs of usage (from the consumers' side) and marginal costs of production (from the supply side) are equivalent. For any standard utility function (monotonically increasing concave or linear), such marginal costs are the sole reason for versioning. Optimality of versioning as a strategy is distribution independent.* ■

Proposition 5 is critical to reconciliation of results from extant models some of whom find versioning of information goods as an optimal monopolist strategy while others find a single product strategy superior. And this proposition is true under both the case of full information and information asymmetry; we shall largely discuss the latter here since it is more general.

For the proofs of Lemma 4 and Proposition 5 we consider a general model where we can now impose specific assumptions on the cost parameters to compare with extant setups. For example if we consider $\lambda = 0$ and maintain the development cost, we get the general setups suggested by Wei and Nault (2008) and Jones and Mendelson (2011) – strictly multiplicative monotonic utility with convex initial development cost and zero marginal (and versioning) costs. We can easily see that there can be no versioning result and a single version will ensue. On the other hand consider a case where $\lambda$ is positive with no initial development costs, we can easily that this leads to a versioning menu but without any distortion in quality for the high types, i.e., all consumers who are served receive their second-best quality. Collectively, these observations tell us that while the capital cost is responsible for quality distortion for the high types, the usage cost is responsible for the versioning decision.

Now consider a physical good equivalent with the traditional multiplicative monotonic utility with convex initial development cost *and* a positive quality-dependent marginal cost, i.e., where $U(\theta) = \theta x - p$ and where the vendor suffers a cost $\lambda x^2$ to serve each consumer. Proposition 5 tells us that this will yield the same versioning strategy as the one for goods with no free disposal. It is interesting to note that even though the vendor suffers the cost in this case, he will offer the same menu thus reconciling with a recent physical goods segmentation result (Hahn 2000). Further examination of the surplus per consumer makes the economic intuition behind this apparent – irrespective of who suffers the cost the vendor maximizes the net surplus per consumer then pays the corresponding rent (to ensure incentive compatibility) and extracts the reminder through price. Thus it does not matter if the loss from consuming a good of a certain quality is through the no free disposal property of the consumer or from the marginal production costs to the vendor.

The above discussion also leads us to carefully examine the objective function given by equation (3). We can see that the only time a menu is derived as a strategy is when there is an interior solution to the term inside the integrand, i.e, the term inside given by equation (8)

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[ \theta x(\theta) - \lambda \left[ x(\theta) \right]^2 - \frac{x(\theta)\left(1 - F(\theta)\right)}{f(\theta)} \right] f(\theta)\, d\theta \tag{8}$$

is concave in $x(\theta)$ . Note that this is always true in the case of goods with no free disposal or when there is a convex marginal cost such as for physical goods. In other words, we can categorically prove that versioning is optimal only in the presence of usage costs or marginal costs of production. Extant utility functions (increasing and multiplicative) considered in information goods literature can never lead to versioning results.

## References

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

Boehm, B., et al. (2000). "Software development cost estimation approaches—A survey." Annals of Software Engineering **10**(1): 177-205.

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

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

Desai, P., et al. (2002). "Product Differentiation and Commonality in Design: Balancing Revenue and Cost Drivers." Management Science **47**(1): 37-51.

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

Hahn, J. H. (2000). Monopoly Quality Differentiation with Top-quality Dependent Fixed Costs. Department of Economics, Keele University.

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

Jones, R. and H. Mendelson (2011). "Information goods vs. industrial goods: Cost structure and competition." Management Science **57**(1): 164-176.

Laffont, J.-J. and D. Martimort (2001). The theory of incentives: the principal-agent model, Princeton University Press.

Mas-Colell, A., et al. (1995). Microeconomic Theory, Oxford University Press, USA.

Milian, M. (2010). "'Junkware' comes standard on Verizon, T-Mobile smart phones." Los Angeles Times: July 15.

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

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

Varian, H. R. (1997). Versioning Information Goods. Internet Publishing and Beyond: Economics of Digital Information and Intellectual Property. B. Kahin and H. R. Varian. Cambridge, MA, MIT Press: 1-13.

Wei, X. D. and B. R. Nault (2008). Vertically differentiated information goods: Monopoly power through versioning, Working paper, University of Calgary, Alberta, Canada.

## 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:

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

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

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

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 $\theta$ can be a solution to $\hat{\theta}$ since that can happen only when $G\left(\theta\right)$ is a linear function of $\theta$. 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:

$$\max_{\{x(\theta),p(\theta)\}} \int_{\underline{\theta}}^{\overline{\theta}} p(\theta) f(\theta) d\theta$$

$$\text{s.t.} \quad U(\theta) \geq 0 \qquad \text{(IR)}$$

$$U(\theta) \geq U_{\tilde{\theta}}(\theta) \qquad \text{(IC)}$$

We first focus on the IC condition. Suppose the vendor offers a quality/feature-price schedule $\{x(\theta), p(\theta)\}$ for every type $\theta$. 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 $\theta$. In other words, the customers' incentive compatibility constraints (ICs) must be satisfied. We represent the utility of a customer of type $\theta$ who declares her type to be $\tilde{\theta}$ as $U_\theta(\tilde{\theta})$. Hence, it must be that:

$$U_\theta(\theta) \geq U_\theta(\tilde{\theta}) \Rightarrow \theta x(\theta) - \lambda x^2(\theta) - p(\theta) \geq \theta x(\tilde{\theta}) - \lambda x^2(\tilde{\theta}) - p(\tilde{\theta}) \tag{9}$$

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 $\theta$ would result in lower utility for her. Corresponding to Equation (5), we get

$$U_{\tilde{\theta}}(\tilde{\theta}) \geq U_{\tilde{\theta}}(\theta) \Rightarrow \tilde{\theta} x(\tilde{\theta}) - \lambda x^2(\tilde{\theta}) - p(\tilde{\theta}) \geq \tilde{\theta} x(\theta) - \lambda x^2(\theta) - p(\theta) \tag{10}$$

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

$$[x(\theta) - x(\tilde{\theta})][\theta - \tilde{\theta}] \geq 0 \tag{11}$$

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

$$x'(\theta) \geq 0 \tag{12}$$

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

$$\theta x'\left(\theta\right) - 2\lambda x\left(\theta\right)x'\left(\theta\right) - p'\left(\theta\right) = 0 \tag{13}$$

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

$$\theta x''\left(\theta\right) - 2\lambda\left[x'^2\left(\theta\right) + x\left(\theta\right)x''\left(\theta\right)\right] - p''\left(\theta\right) < 0 \tag{14}$$

Differentiating Equation (9) with respect to $\theta$, we get

$$x'\left(\theta\right) + \theta x''\left(\theta\right) - 2\lambda\left[x'^2\left(\theta\right) + x\left(\theta\right)x''\left(\theta\right)\right] - p''\left(\theta\right) = 0 \tag{15}$$

Substituting from Equation (11) in (10) we obtain $x'\left(\theta\right) \geq 0$. 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 $\left(\frac{\partial^2 U\left(x,p,\theta\right)}{\partial x\partial\theta} = \frac{\partial\left(\theta - 2\lambda x\right)}{\partial\theta} = 1\right)$ 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:

$$U\left(\theta\right) = \theta x\left(\theta\right) - \lambda x^2\left(\theta\right) - p\left(\theta\right) \tag{16}$$

Differentiating both sides of the above equation with respect to $\theta$, we get:

$$U'\left(\theta\right) = x\left(\theta\right) + \theta x'\left(\theta\right) - 2\lambda x\left(\theta\right)x'\left(\theta\right) - p'\left(\theta\right) \tag{17}$$

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

$$U'(\theta) = x(\theta) \tag{18}$$

Integrating Equation (14) between the limits $\underline{\theta}$ and $\theta$, 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

$$U(\theta) = \int_{\underline{\theta}}^{\theta} x(y)dy \tag{19}$$

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

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\theta x(\theta) - \lambda x^2(\theta)\right] f(\theta)d\theta - \int_{\underline{\theta}}^{\bar{\theta}} \left[\int_{\underline{\theta}}^{\theta} x(y)\,dy\right] f(\theta)d\theta \tag{20}$$

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

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\theta - \lambda x(\theta) - \frac{1 - F(\theta)}{f(\theta)}\right] x(\theta)f(\theta)d\theta \tag{21}$$

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

$$x^*\left(\theta\right) = \frac{\theta - \left[\dfrac{1-F\left(\theta\right)}{f\left(\theta\right)}\right]}{2\lambda} \tag{22}$$

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 $\left[\dfrac{1-F\left(\theta\right)}{f\left(\theta\right)}\right]$ is decreasing in $\theta$. Hence $x^*\left(\theta\right)$ is increasing in $\theta$ 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 $\theta_L^*$. Then we have:

$$\theta - \left[\frac{1-F\left(\theta\right)}{f\left(\theta\right)}\right] = 0 \tag{23}$$

The solution to the above equation, $\theta_L^*$, 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 $\theta_H^*$. This point is the solution to:

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

$$\text{or, } \theta - \frac{1-F\left(\theta\right)}{f\left(\theta\right)} - 2\lambda x_H = 0 \tag{24}$$

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

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

This can be simplified to:

$$\theta_H^* \left[1 - F\left(\theta_H^*\right)\right] = 2x_H^* \left[\lambda\left[1 - F\left(\theta_H^*\right)\right] + c\right] \tag{25}$$

Substituting $\theta_H^*$ in place of $\theta$ in Equation (20) and solving it simultaneously with Equation (21) we obtain $\theta_H^*$ 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 $O_C(x)$. Hence, we have:

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

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

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

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

$$\left.\frac{dO_C(x)}{dx}\right|_{x=x_H^*} = \int_{\hat{\theta}}^{\theta_H^*} \left[\theta - \hat{\theta}\right] f(\theta)\, d\theta + \int_{\theta_H^*}^{\bar{\theta}} \left[1 - F(\theta)\right] d\theta \tag{26}$$

Note that $\hat{\theta}^*(x_H^*) = 2\lambda x_H^*$ and from Lemma 2, $\theta_H^*(x_H^*) = 2\lambda x_H^* + \frac{1 - F\left(\theta_H^*\left(x_H^*\right)\right)}{f\left(\theta_H^*\left(x_H^*\right)\right)}$. 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 $\theta \leq \overline{\theta}$. Thus, we have shown that $\left.\frac{dO_C(x)}{dx}\right|_{x=x_H^*} > 0$. Further, we know that $\left.\frac{dO_C(x)}{dx}\right|_{x=\hat{x}^*} = 0$ 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 $x^*\left(\theta\right) = \frac{\theta}{2\lambda}$ and from Proposition 2, the quality served under incomplete information is $x^*\left(\theta\right) = \frac{\theta - \frac{1 - F(\theta)}{f(\theta)}}{2\lambda}$. Since $F\left(\theta\right) < 1$, 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 $\theta$). Further, $\hat{x}^* > x_H^*$. Hence a reduced quality is served under information asymmetry.

Suppose $\hat{\theta}^* > \theta_H^*$. The customer indexed by $\theta_H^*$ 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 $\theta$ 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 $\theta$). 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 $\lambda$, 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* $\theta_L^*$

From Lemma 2, it is easy to see that $\theta_L^*$ does not depend on either $c$ or $\lambda$.

*Comparative Statics of* $\theta_H^*$

$\theta_H^*$ 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 $\lambda$ increases, the solution to the above equation, i.e., $\theta_H^*$ increases. Also, as $c$ increases, $x_H^*$ reduces and hence $\theta_H^*$ reduces.

**Proof of Lemma 4 and Proposition 5**

We are able to show that Lemma 4 and Proposition 5 are applicable to a more general utility function (and therefore applicable to our functional form as well). So we shall use new notation for this proof along.

Let consumers of type $\theta$ enjoy a utility $u(\theta, x) - p$ when choosing some quality $x$ and paying a monetary transfer $p$. We assume that the standard sorting condition, $u_{\theta x}(\theta, x) > 0$ holds for positive $x$ with $u(0, x) > 0$ and $u(\theta, 0) = 0$.

Providing quality $x$ is costly to the monopolistic firm with a marginal cost $c(x)$ with $c(0) = 0$. The goal of the monopolistic firm is to design a price schedule $\{x(\theta), p(\theta)\}$ to maximize its profit. However, the monopolistic firm cannot arbitrarily choose $\{x(\theta), p(\theta)\}$ as the consumer's ability to choose $x$ must be respected (self-selection condition). Incentive compatibility requires that $x(\theta)$ weakly increases in $\theta$ $\left(x'(\theta) \geq 0\right)$ and the consumer's marginal surplus is $v'(\theta) = u_\theta(\theta, x(\theta))$ as given by the envelope theorem (with inequality constraints)). The payoff per consumer can be written as the valuation per consumer less the consumer surplus and the marginal cost

$$R(\theta) = u(\theta, x(\theta)) - v(\theta) - c(x(\theta)) \tag{27}$$

So the total expected profits for the firm is given by

$$\begin{aligned} &\max_{\{x'(\theta)\}} \int_{\underline{\theta}}^{\bar{\theta}} R(\theta) f(\theta) d\theta \\ &\text{s.t.} \quad x(\theta) \le x_H, x'(\theta) \ge 0 \\ &\qquad v'(\theta) = u_\theta(\theta, x(\theta)) \end{aligned} \tag{28}$$

Since the sorting condition, $u_{\theta x}(\theta, x) > 0$ holds, and $u(\theta, 0) = 0$, for any $x > 0$, we have

$$v'(\theta) = u_\theta(\theta, x(\theta)) - u_\theta(\theta, 0) > 0 \tag{29}$$

Therefore the individual rationality constraint is satisfied everywhere if it holds at the lowest type.

Hence $v(\theta) = \int_{\underline{\theta}}^{\theta} u_\theta(t, x(t)) dt$. Standard transformation yields

$$\begin{aligned} &\max_{x(\theta)} \int_{\underline{\theta}}^{\bar{\theta}} \left( u(\theta, x(\theta)) - \frac{1 - F(\theta)}{f(\theta)} u_\theta(\theta, x(\theta)) - c(x(\theta)) \right) f(\theta) d\theta \\ &\text{s.t.} \quad x'(\theta) \ge 0 \\ &\qquad x(\theta) \le x_H \end{aligned} \tag{30}$$

The firm essentially needs to find $x$ that maximizes the surplus extracted from each consumer of type $\theta$, i.e., we can easily employ point-wise maximization to the term inside the integrand in equation (30).

Our aim is to show that versioning can be the optimal strategy only if

$$R_{xx}(\theta, x) < 0 \quad \text{for some } \theta \in [\underline{\theta}, \bar{\theta}] \tag{31}$$

In other words, the marginal rent to the firm needs to be strictly concave for some $\theta \in [\underline{\theta}, \bar{\theta}]$ to make versioning an optimal monopolist strategy.

Suppose that the payoff function satisfies $R_{xx}(\theta, x) = 0$ for all $\theta \in [\underline{\theta}, \bar{\theta}]$, it implies that the first order derivative of the firm's payoff function $u_x(\theta, x) - \frac{1 - F(\theta)}{f(\theta)} u_{\theta x}(\theta, x) - c_x(x)$ is invariant with respect to $x$. Then the firm will simply assign $x_H$ to consumers of type $\theta$ with

$u_x(\theta,x)-\frac{1-F(\theta)}{f(\theta)}u_{\theta x}(\theta,x)-c_x(x)\geq 0$, whereas assigning zero quality to consumers of type $\theta$ with $u_x(\theta,x)-\frac{1-F(\theta)}{f(\theta)}u_{\theta x}(\theta,x)-c_x(x)<0$.

Given $u_{\theta x}(\theta,x)>0$ and $u_{\theta\theta x}(\theta,x)=0$ and the monotone hazard rate, we have

$$R_{\theta x}(\theta,x)=\frac{\partial}{\partial\theta}\left[u_x(\theta,x)-\frac{1-F(\theta)}{f(\theta)}u_{\theta x}(\theta,x)-c_x(x)\right]=\left(1-\frac{d}{d\theta}\left(\frac{1-F(\theta)}{f(\theta)}\right)\right)u_{\theta x}(\theta,x)>0 \qquad (32)$$

Because the marginal rent is increasing in $\theta$ for all values of $x$, it must be unique if there exists some $\theta$ such that $u_x(\theta,x)-\frac{1-F(\theta)}{f(\theta)}u_{\theta x}(\theta,x)=0$. Denote this unique point by $\hat{\theta}$. Then $x(\theta)=x_H$ for all $\theta\in[\hat{\theta},\overline{\theta}]$ and $x(\theta)=0$ otherwise. It implies that, if there is a marginal type $\hat{\theta}$ of consumers the firm finds profitable to serve, then all consumers with $\theta$ larger than $\hat{\theta}$ will be served with quality $x_H$. This is the optimal schedule with $R_{xx}(\theta,x)=0$ for all $\theta\in[\underline{\theta},\overline{\theta}]$.

Since versioning by definition requires that there exist at least two different positive qualities, the proof above shows that versioning is not the optimal strategy if $R_{xx}(\theta,x)=0$ for all $\theta\in[\underline{\theta},\overline{\theta}]$.

Since $R(.)$ is defined by equation (27), equation (31) can be true if

(a) Consumer utility function itself is strictly concave, i.e., $u_{xx}(\theta,x)<0$ or/and

(b) if $u(\theta,x)$ is monotonically increasing concave or linear, the marginal cost $c(x)$ suffered by the firm is strictly convex i.e., $c_{xx}(x)>0$.

From a monopolist' point of view (a) and (b) are duals of each other since the net surplus that can be extracted is the same.