Read-only Versions for Free and for Profit: Functional Quality Differentiation Strategies of a Software Producing Monopoly*

Gergely Csorba†

Central European University, Budapest

2002-03-27

Abstract

This paper analyzes a special quality discrimination policy of a software firm, which removes some functions of his product and sells the damaged version at a lower price. Because of the network externality exerted by the users, the firm may find it profitable to distribute the degraded version at zero price, these are the so called free read-only versions.

The paper derives the conditions for introducing the read-only version, and the conditions for selling it for free with a continuum type of consumers. Welfare consequences are also examined, and it is shown that functional versioning may result in a Pareto-improvement.

Keywords: network externalities, multiple equilibria, quality differentiation, sorting condition

JEL Classification: D42, L12, L15, L86

*The main part of this paper was written when the author spent a year at the University of Toulouse. The author is very much indebted to Bruno Jullien, who always encouraged the further development of the model, and thanks to Patrick Rey, Jean Tirole and participants of the microeconomics workshop at CEU for comments.

†Postal address: Central European University, Economics Department, Nador u. 9, 1051 Budapest, Hungary. E-mail address: cphcsg01@phd.ceu.hu

1 Introduction

Different versions of the same software are frequently encountered on the software market . In this paper, different versions will mean different forms of a software of one generation, i.e. when versions from one date of release differ in the functions they provide.[1] Producing different version of a software is of small difficulty: after the software firm has developed the so called "top product", he simply removes some of its functions, and basically both the degradation and reproductions costs after are negligible. The typical versioning policy we may observe is the production of the so called read-only versions of a software, when the firm removes the writing function of the full version, and sells the degraded version at a lower price, often for free, making it downloadable from the Internet as well.

Several examples can be found, just to give the more closest: this paper has been written in Scientific Workplace, and Mathematica has been used for some calculations - both software have a full and a read-only version. Finally it has been formatted into PDF, the file format developed for the archetypes of this phenomenon - the Adobe Acrobat remains quite expensive, and the Acrobat Reader is legally downloadable from a thousand webpages and can be found in most of the Internet subscription packages (in the latter case users do not face even download costs).[2] Even Microsoft has (free) read-only versions for his Office programs.[3]

Mussa and Rosen [1978] and Maskin and Riley [1984] are the classical references to present quality discrimination strategies of a monopoly. However, these works treated goods of different quality independently, and moreover, they assumed the cost of producing a good of higher quality to be increasing. Deneckere and McAfee [1996] were the first to analyze the situation from other way around - the good of the highest quality is developed first, and then the low quality one can be produced by degrading the first on additional costs. Although their seminal

[1]Loosely speaking, they differ in their version name, not in their version number.

[2]The price of Adobe Acrobat 5.0 is 249$. Adobe Systems Inc. is still following this strategy, although there exist other software that can produce PDF format, which are less expensive or shareware. An explanation for this can be the large installed base of Adobe and the switching costs users are facing.

[3]For other examples, consult chapter 3 of Shapiro and Varian [1999].

model of damaged goods can be more properly applied for hardware products, the general idea remains valid for the software market as well: "By producing an inferior substitute, the manufacturer can sell to customers who do not value the superior product so much, without decreasing demand for the superior product very much" (Deneckere and McAfee [1996], p. 1). Additionally, the authors show that this degradation policy may lead to a strict Pareto-improvement, i.e. all the consumers and the monopoly benefit from quality discrimination of this type.

Jing [2000] extends the analysis of software versioning strategies in the presence of network externalities. He assumes that consumers have a stand-alone utility from using a software, plus they enjoy some positive externalities exerted by other users of the same software. In this model versioning occurs since the low version enlarges market size, and the increased installed base makes the product more valuable, so consumers are eager to pay an even higher price for it. Without network externalities, only one version would be produced. An interesting result of the paper is that the monopoly develops exactly two versions, because a "middle quality" would always decrease profit.[4] Conditions for the free degraded product to exist are also derived.

However, both the Deneckere-McAfee and the Jing model were models of proportional quality degradation: quality was measured by an universal number, by which each consumer's utility was multiplied. Hahn [2001] assumes another type of preferences, which may fit the nature of software more appropriately. In his model of Lancesterian style a software is a bundle of functions it can perform, and consumers derive different utilities by using these different functions. Hahn considers two basic functions of a software, writing and reading, the original product being the one that can do both. The firm has to decide whether to unbundle this product and sell a mix of the bundle (the full version) and one of its components (the read-only version) simultaneously or only the bundle - this is the opposite direction of the standard bundling models.[5]

Let us consider the following example to illustrate this phenomenon.[6] Suppose

[4]However, examples on why "goldilocks pricing" can be beneficial in some cases are found in chapter 3 of Shapiro–Varian [1999].

[5]The question of mixed bundling was first analyzed by Adams and Yellen [1976]. There are some recent papers concerning bundling in the presence of network externalities, see Bakos and Brynjofsson [1999, 2000].

[6]This example was suggested by Bruno Jullien.

there are two potential groups of users of a text processor: members of the first group are interested in writing scientific papers, and the more people read their work, the happier they are (e.g. professors publishing their work on their website); the second group has no writing ambitions at all (e.g. students). On top of that, suppose that no user likes reading at all, but they read every paper that is published (professors because of their colleagues, students because of the exams). When there is just one version of the software, it includes the writing and the reading function, and it is sold to the first group. However, introducing the read-only version, and selling it to the second group would prove profitable, despite the facts that it has to be distributed at a zero price, and that the targeted group will not benefit from using it. The increased reading base positively affects the benefits derived by the first group, hence the software firm could set a higher price for them, increasing thereby his profit.

Hahn shows in a simple model of two types of consumers that this functional degradation strategy may prove very effective in screening consumers, especially if consumer's valuation for the different functions are (favorably) negatively correlated. The read-only version may be introduced even under cost, because the loss the monopoly has to bear can be compensated by reaping the increased surplus of the full version users (cross-subsidization). Versioning may result in a strict Pareto-improvement as well, if it increases the amount of software sold.

The contribution of this paper is to generalize Hahn's model in a continuous type framework. Moving into a more complicated setup, we have to solve first the usual problem of models with network externalities, that consumers face a coordination problem in their decisions, which may generate multiple equilibria.[7] However, by assuming positive correlation of the consumers' valuations for the two functions, we can use the Pareto-criterion to select one of these equilibria and reformulate the monopoly's decision problem to handle the problem in a more practical way. Then necessary and sufficient conditions similar to the sorting condition are derived to characterize the situations where the firm finds it profitable to introduce the read-only version and where it is provided for free. It is shown that versioning enlarges the size of the market, although it always decreases the number of the full version sold. The welfare implication of versioning strategies do not give unambiguous results; however, some sufficient conditions are derived

[7] Farrell and Saloner [1985] and Katz and Shapiro [1985] were the first to raise this problem.

when it results in a Pareto-improvement.

The rest of the paper is organized as follows. Section 2 sets up the model and discusses briefly its assumptions, then Section 3 examines the benchmark case of selling the full version only (no versioning). Section 4 derives the necessary and sufficient conditions for versioning to occur and for the read-only version to be sold for free as well. Section 5 deals with the welfare effects of introducing the read-only version. Section 6 summarizes and discusses the possible extensions of the model. The more technical proofs are delegated to the Appendix.

2 The model

Take the case of a software firm who is the sole supplier of a two-way functional software: the read function allows users to read products made (and freely distributed) by the users owning a software of writing function. Naturally, the software with writing function includes the reading function as well, so the monopolist can sell (at most) two types of product: the full (write and read) and the read-only versions. Assume that once the firm developed the full and possibly the read-only version, he faces no reproduction and selling costs.

Each consumer buys at most one unit of good (the full version or the read-only), at the price of p_f or p_r, respectively. The utility of each consumer is assumed, as in Hahn [2001], to be additively separable in the two valuations for the functions of the software and the price, and is the following for the consumer of type i:

$$U_i = \left\{ \begin{array}{l} (n_f + n_r)v_i^w + n_f v_i^r - p_f, \text{ if he buys the full version,} \\ n_f v_i^r - p_r, \text{ if he buys the read-only version,} \end{array} \right.$$

where $v_i^w$and v_i^r stand for valuations for the writing and reading function, n_f and n_r are the number of consumers who buy the full and the read-only versions. So the good is assumed to be a pure network good, it has a value only when it is used also by other consumers. If a user owns the full version, he enjoys the network externality exerted by the consumers able to read him (or more precisely, his works), plus the network externality exerted by the consumers he reads. Owners of the read-only software benefit only from the latter externality. The consumers' type is not observable by the firm (or even if it is, he cannot discriminate among consumers).

Now index the consumers by their valuation for the writing function, let this index be v. We assume that v has a continuous cumulative distribution function $F(v)$, with a density function $f(v) > 0$ normalized to the support $[0, 1]$. We should employ the necessary assumption familiar in adverse selection models, namely the monotone hazard rate property: $\frac{d}{dv}\frac{F(v)}{f(v)} > 0$, which will be provided by the convexity of the distribution function: $f'(v) \geq 0$.[8] A consequence of the monotone hazard rate property, which we will use extensively in the following analysis, is that $\frac{1-F(v)}{f(v)}$ is a decreasing function of v.

We assume that a consumer with higher valuation for the writing function has higher valuation for the reading function as well (positive correlation), and there exists a continuous function $g(v), g'(v) > 0$, which gives the valuation for the reading function of a v-type consumer. Without the loss of generality we can normalize the valuations such that consumers having zero valuation for the writing function have zero valuation for the reading function:$g(0) = 0$. We also assume that the valuation for the reading function changes more slowly as the valuation for the writing function increases, i.e. $g(v)$ is a concave function of v: $g''(v) \leq 0$, and that the valuation for the reading function does not extend the valuation for the writing function: $g(1) \leq 1$.

Consumers are assumed to form expectations about the size of n_f and n_r, and these expectations must be fulfilled in equilibrium (rational expectations). On top of that we also assume that if at a certain pair of price there exist more than one pair of (n_f, n_r) that satisfies this equilibrium condition and they can be Pareto-ranked (i.e. one of them makes everyone better off than the others), they expect this allocation to prevail in equilibrium. This is the so called Pareto criterion.[9]

The timing of the model is the following. First, the monopoly decides whether to introduce the read-only version or not, and sets the price(s) of his product(s). Second, consumers observe the prices, and form their (rational) expectations about the expected size of the market, (n_f^e, n_r^e). Third, each consumer decides which good to purchase or buys nothing.

[8]The latter assumption is needed only for ensuring that the second-order conditions of the problem will be be always satisfied. In fact, it would be sufficient to state a condition ensuring that the distribution function is "not too concave", but it would entail more technical difficulties.

[9]On the use of this criterion in the presence of network externalities, see Ellison and Fudenberg [2000].

3 Selling the full version only

Let us examine first the case of the monopoly producing only the full version. When the monopoly sets a price p_f for his product, consumer v_f is indifferent between buying and not buying the product when

$$n_f^e[v_f + g(v_f)] - p_f = 0,$$

where n_f^e indicates the expected value of consumers buying the good. Since the valuation for the full version is an increasing function of the type, all the consumers having a higher type than v_f will also purchase the good, and no one having a lower type than v_f will buy it. Because consumers form rational expectations, in equilibrium we have to have that $n_f^e = 1 - F(v_f)$, which means that

$$p_f = [1 - F(v_f)][v_f + g(v_f)]. \tag{1}$$

It can be easily checked that the assumptions made on $F(\cdot)$ and $g(\cdot)$ ensure that $p_f(v_f)$ is a strictly concave function. Because $p_f(0) = p_f(1) = 0$, for each price there exist more than one v_f (so more than one quantity) that satisfies equilibrium condition (1), as usual in the presence of network externalities. But note that the smaller v_f (the higher quantity) that satisfies equation (1) gives higher value to the network good, so everyone would better off by coordinating on this quantity. So as a consequence of the Pareto criterion, for each price p_f the monopoly chooses there exist only one v_f the consumers expect, which will be fulfilled in equilibrium.

Having derived a demand function, we conjecture that the firm could write down and solve his maximization problem in respect of the marginal consumer's type instead of the price. This equivalence is confirmed by the following lemma.

Lemma 1 *It is equivalent to express and maximize the firm's profit as function of the price, $\Pi(p_f)$, and as function of the marginal consumer's type, $\Pi(v_f)$.*

Proof: See Appendix. ■

The monopoly can thus maximize his profit in v_f, and then he will charge the price p_f that satisfies equilibrium condition (1). At this price the expected (and fulfilling) quantity will be the profit-maximizing one.

The firm's profit can be thus written as

$$\Pi(v_f) = p_f(v_f)[1 - F(v_f)] = [1 - F(v_f)]^2[v_f + g(v_f)].$$

Maximizing in v_f gives the following first-order condition:[10]

$$\frac{1 - F(v_f)}{f(v_f)} = \frac{2[v_f + g(v_f)]}{1 + g'(v_f)}. \tag{2}$$

Equation (2) always gives a unique solution for v_f, since the left-hand side is decreasing in v_f by the monotone hazard rate property and takes all the value between 0 and 1, while the right-hand side is increasing in v_f (the numerator is increasing, the denominator is non-increasing in v_f) and takes the value of 0 at $v_f = 0$.

4 Introducing the read-only version

Now let us turn to the case when the read-only version is introduced. If the monopoly sets prices p_w and p_r for the full and read-only versions, a consumer of type v purchases the full version if he derives non-negative utility by its use and has no incentive to switch to the other version:[11]

$$(n_f^e + n_r^e)v + n_f^e g(v) - p_w \geq n_f^e g(v) - p_r, \text{ and} \tag{3}$$

$$(n_f^e + n_r^e)v + n_f^e g(v) - p_w \geq 0. \tag{4}$$

Similarly, a consumer of type v' buys the read-only version, if

$$n_f^e g(v') - p_r \geq (n_f^e + n_r^e)v' + n_f^e g(v') - p_w, \text{ and} \tag{5}$$

$$n_f^e g(v') - p_r \geq 0. \tag{6}$$

Let us denote by v_w the type of the consumer who is indifferent between purchasing the two versions and by v_r the type of the consumer who is indifferent

[10] In the optimum $0 < v_f < 1$ (other values give non-positive profits), so we can use unconstrained optimisation. We can ignore second-order conditions as well, since we will see in the next subchapter that this maximization program is part of a more general one, for which the first-order conditions are necessary and sufficient.

[11] Because of the continuum of consumers, an individual's contribution to the network is negligible, so individual switching leaves the number of full and read-only version users unchanged.

between purchasing the read-only version and not purchasing at all, so

$$
\begin{aligned}
(n_f^e + n_r^e)v_w + n_f^e g(v_w) - p_w &= n_f^e g(v_w) - p_r, \text{ and} \\
n_f^e g(v_r) - p_r &= 0.
\end{aligned}
$$

Rearranging gives the following expressions for the prices:

$$p_w - p_r = (n_f^e + n_r^e)v_w, \tag{7}$$

$$p_r = n_f^e g(v_r), \tag{8}$$

$$p_w = (n_f^e + n_r^e)v_w + n_f^e g(v_r). \tag{9}$$

The first two equations show the intuition we expected: the price for the marginal read-only version buyer is equal to his valuation of "reading everybody who can write", while the price difference between the full and read-only versions for the marginal full version buyer is equal to his valuation of "writing to everybody who can read".

We can also see easily that the marginal consumers' types "cut" the mass of consumers into different purchasing groups.

Lemma 2 *Every consumer having a higher type than v_w purchase the full version, consumers of type $v_r < v < v_w$ buy the read-only version, and consumers with a type less than v_r buy nothing.*

Proof: See Appendix. ■

Consumers form rational expectations, so in equilibrium the expectations about the quantities purchased have to be fulfilled: $n_f^e = 1 - F(v_w), n_r^e = F(v_w) - F(v_r)$. By rewriting equations (8) and (9) we have the following two equilibrium conditions:

$$p_w = [1 - F(v_r)]v_w + [1 - F(v_w)]g(v_r), \text{ and} \tag{10}$$

$$p_r = [1 - F(v_w)]g(v_r). \tag{11}$$

As seen before, there can be multiple pairs of (v_w, v_r) that satisfy these equilibrium conditions at given prices (p_w, p_r). But if we compare two different equilibrium pairs, we see that they can be ordered:

Lemma 3 *Let* (v_w, v_r) *and* (v'_w, v'_r) *be two equilibrium pairs that satisfy equations (10) and (11) at a given pair of prices, such that* $v_w < v'_w$*. Then* $v_r < v'_r$ *has to be satisfied.*

Proof: $[1 - F(v_w)]g(v_r) = [1 - F(v'_w)]g(v'_r)$*, since both pairs satisfy equation (11). If* $v_w < v'_w$*, then* $1 - F(v_w) > 1 - F(v'_w)$*, which implies* $v_r < v'_r$*.* ■

Because of the positive externalities in both quantities, the smaller pair gives higher utility to every consumer, so by the Pareto-principle they expect the smallest pair (v_w, v_r) satisfying equations (10) and (11) at given prices (p_w, p_r) to be the equilibrium.

Having now a demand function $h(p_w, p_r)$ we can follow the same way used in the previous section to solve the monopoly's problem by writing down and maximizing his profit as a function of the marginal consumers' type..

Lemma 4 *It is equivalent to express and maximize the firm's profit as the function of the prices,* $\Pi(p_w, p_r)$*, and as the function of the marginal consumers' type,* $\Pi(v_w, v_r)$*.*

Proof: See Appendix. ■

So after finding the profit-maximizing v_w and v_r, the firm will charge the prices p_w and p_r that satisfies equilibrium condition (10) and (11). At these prices the expected (and fulfilling) quantities will be the profit-maximizing ones.

After using equations (10) and (11), the firm's profit can be simplified to the following form:

$$\begin{aligned} \Pi(v_w, v_r) &= [1 - F(v_w)]p_w + [F(v_w) - F(v_r)]p_r = \\ &= [1 - F(v_w)][1 - F(v_r)][v_w + g(v_r)]. \end{aligned}$$

The monopoly's profitmaximisation problem can be then written as:

$$\begin{aligned} \underset{\{v_w, v_r\}}{Max}\ \Pi(v_w, v_r) &= [1 - F(v_w)][1 - F(v_r)][v_w + g(v_r)] \\ s.t.\ 0 &\leq v_r \\ v_r &\leq v_w \\ v_w &\leq 1. \end{aligned}$$

The third constraint is always slack in optimum, since $v_w \geq 1$ would mean no selling of the full version, so the read-only version would lose its value and no positive profit could be made. Let us denote by λ and μ the Kuhn-Tucker multipliers of the first and the second constraint, respectively.

The marginal types' marginal impact on the profit are the following:

$$\frac{\partial \Pi}{\partial v_w} = [1 - F(v_r)][1 - F(v_w) - f(v_w)(v_w + g(v_r))], \text{ and} \tag{12}$$

$$\frac{\partial \Pi}{\partial v_r} = [1 - F(v_w)][(1 - F(v_r))g'(v_r) - f(v_r)(v_w + g(v_r))]. \tag{13}$$

After some rearrangement the first-order conditions of the firm's profitmaximisation problem are the following:

$$\frac{1 - F(v_w)}{f(v_w)} + \frac{\mu}{[1 - F(v_r)]f(v_w)} = v_w + g(v_r), \text{ and} \tag{14}$$

$$\frac{1 - F(v_r)}{f(v_r)} g'(v_r) + \frac{\lambda - \mu}{[1 - F(v_w)]f(v_r)} = v_w + g(v_r). \tag{15}$$

It can be shown that the profit $\Pi(v_w, v_r)$ function possesses a favorable analytical property, which will prove useful in the following analysis.

Lemma 5 *The function $\Pi(v_w, v_r)$ is strictly quasi-concave.*

Proof: See Appendix. ■

The strict quasi-concavity of the function $\Pi(v_w, v_r)$ facilitates considerably our work. First, the set of maximizers of $\Pi(v_w, v_r)$ is a singleton, so the first-order conditions define a unique solution. Moreover, the first-order conditions are necessary and sufficient for the optimum if one of the constraints is binding, so we have to check the second-order conditions only when they are all slack.[12]

Depending on which of the two constraints is binding we can distinguish three cases, which we develop in the following subsections.[13] In subsection 4.4 we present an example, which illustrates how the changes in the demand parameters result in the different versioning cases.

[12] The second order conditions will be satisfied in this case, see Appendix.

[13] The two constraints cannot be binding simultaneously in optimum, since then $v_w = v_r = 0$ and $\Pi(v_w, v_r) = 0$.

4.1 The non-versioning case

When the second constraint is binding, $v_w = v_r$ and there is no versioning. The first constraint must be then slack, so $\lambda = 0$. Rewriting the first-order conditions (14) and (15) we have

$$\begin{aligned} \frac{1-F(v_w)}{f(v_w)} + \frac{\mu}{[1-F(v_w)]f(v_w)} &= v_w + g(v_w), \text{ and} \\ \frac{1-F(v_w)}{f(v_w)} g'(v_w) - \frac{\mu}{[1-F(v_w)]f(v_w)} &= v_w + g(v_w). \end{aligned}$$

Adding these two constraints yields exactly condition (2), which was the first-order condition of selling the full version only, and this is indeed the case. So $v_w = v_r = v_f$. Expressing μ from these two equations we have

$$\begin{aligned} \mu &= -[1-F(v_f)][1-F(v_f) - f(v_f)(v_f + g(v_f))] = \\ &= [1-F(v_f)][(1-F(v_f))g'(v_f) - f(v_f)(v_f + g(v_f))]. \end{aligned}$$

Comparing this expression with equations (12) and (13) we see that

$$\mu = -\frac{\partial \Pi(v_f, v_f)}{\partial v_w} = \frac{\partial \Pi(v_f, v_f)}{\partial v_r}.$$

For μ to be non-negative, we thus have to have that $\frac{\partial \Pi(v_f,v_f)}{\partial v_w} \leq 0$ and $\frac{\partial \Pi(v_f,v_f)}{\partial v_r} \geq 0$, so when being in the optimum by selling the full version only, the firm has no incentive to choose a larger v_w or a smaller v_r, i.e. to discriminate. Since $\Pi(v_w, v_r)$ is strictly quasi-concave, this local maximum is a global maximum as well, and because one of the constraints is slack, these conditions are also sufficient for this case.

Proposition 6 *If $v_w = v_r = v_f$, then $g'(v_f) \geq 1$.*

Proof: For μ to be non-negative, $1 - F(v_f) - f(v_f)(v_f + g(v_f))$ has to be non-positive, and $(1 - F(v_f))g'(v_f) - f(v_f)(v_f + g(v_f))$ has to be non-negative. Rearranging the terms we have that

$$\begin{aligned} \frac{1-F(v_f)}{f(v_f)} &\leq v_f + g(v_f), \ \textit{and} \\ \frac{1-F(v_f)}{f(v_f)} g'(v_f) &\geq v_f + g(v_f). \end{aligned}$$

These equations can be satisfied only if $g'(v_f) \geq 1$. ■

Since the first-order conditions give necessary and sufficient conditions for the optimum in this case, the previous proposition works in the reverse direction as well.

Corollary 7 *The firm chooses not to introduce the read-only version if and only if $g'(v_f) \geq 1$. If $g'(v_f) < 1$, the firm introduces the read-only version.*

This condition bears some similarity with the sorting condition familiar in adverse selection models (or more precisely with the case when it is not satisfied). If $g'(v_f) \geq 1$, in v_f the valuation for the reading function changes more rapidly than the valuation for the writing function, so any attempt for screening the types would only harm the profit of the firm. Versioning would decrease the externality created by writing (which are enjoyed by reading), and although it creates some externality by extending the reading base, the loss would be bigger, since the number of readers is bigger than that of the writers, and the change in the valuation of the "marginal reader" is bigger. So the firm could not increase his profit by versioning and sells only the full version.

4.2 "Standard" versioning

Consider now the case when neither of the constraints is binding, so $\lambda = \mu = 0$. In this case the monopoly decides to introduce the read-only version and sells it at a positive price, because $p_r = [1 - F(v_w)]g(v_r) > 0$, since $0 < v_r < v_w < 1$. By rewriting the first-order conditions (14) and (15) we have that

$$\frac{1 - F(v_w)}{f(v_w)} = v_w + g(v_r), \text{ and} \tag{16}$$

$$\frac{1 - F(v_r)}{f(v_r)} g'(v_r) = v_w + g(v_r). \tag{17}$$

The following lemma will prove useful in the further analysis. Moreover, together with Proposition 9 it strengthens the arguments of the previous subchapter concerning our sorting condition: if $g'(v_r) < 1$, then $g'(v_f) < 1$ by the concavity of $g(v)$, so the firm chooses versioning, which is indeed the case.

Lemma 8 *If $0 < v_r < v_w < 1$, then $g'(v_r) < 1$.*

Proof: Since $v_r < v_w$, by the monotone hazard rate property we have that $\frac{1-F(v_w)}{f(v_w)} < \frac{1-F(v_r)}{f(v_r)}$. The right-hand side of conditions (16) and (17) are the same, *so* $\frac{1-F(v_w)}{f(v_w)} = \frac{1-F(v_r)}{f(v_r)} g'(v_r)$. *For this equation to hold in optimum,* $g'(v_r)$ *must be less than one.* ■

We will now show that the monopoly sells to more consumers by versioning , but sells less quantity of the full version than by selling the full version only.

Proposition 9 *If* $0 < v_r < v_w < 1$*, then* $v_r < v_f < v_w$.

Proof: The proof is by contradiction. First let us suppose that $v_r < v_w \leq v_f$. *Monotone hazard rate property implies* $\frac{1-F(v_f)}{f(v_f)} \leq \frac{1-F(v_w)}{f(v_w)}$. *Replacing each side by using optimum conditions (2) and (16) we have that*

$$\frac{2}{1+g'(v_f)}[v_f + g(v_f)] \leq v_w + g(v_r).$$

But this is a contradiction, since by the concavity of $g(v)$ *and Lemma 8,* $1 + g'(v_f) < 1 + g'(v_r) < 2$*, and* $v_f + g(v_f) \geq v_w + g(v_r)$.

Suppose now that $v_w > v_r \geq v_f$. *Monotone hazard rate property implies* $\frac{1-F(v_r)}{f(v_r)} \leq \frac{1-F(v_f)}{f(v_f)}$. *After replacing each side by using optimum conditions (2) and (17), by rearrangement we have that*

$$v_w + g(v_r) \leq \frac{2g'(v_r)}{1+g'(v_f)}[v_f + g(v_f)].$$

This is a contradiction again, since $v_w + g(v_r) \geq v_f + g(v_f)$*, and* $1 + g'(v_f) \geq 1 + g'(v_r) > 2g'(v_r)$ *by the concavity of* $g(v)$ *and Lemma 8.* ■

The comparison of the prices does not give so unambiguous results. We can easily see, that the price of the read-only version: $p_r = [1 - F(v_w)]g(v_r)$ is smaller than the price of the full version, when it is the only version sold: $p_f = [1-F(v_f)][v_f+g(v_f)]$, since $1-F(v_w) < 1-F(v_f)$ and $g(v_r) < v_f+g(v_f)$ by Proposition 9. But we cannot make such a clear ordering between p_f and the price of the full version in the versioning case: $p_w = [1 - F(v_r)]v_w + [1 - F(v_w)]g(v_r)$.

However, even if we knew which price is bigger, we could not derive direct welfare conclusions from the fact that "versioning leads to a decrease in the price of the full version", for example. If the firm chooses to introduce the read-only version, some consumers who had the full version "before", will purchase the read-only version now, which destroys some externalities, while on the other side the firm attracts new readers by versioning , which creates new externalities. So if a consumer still purchases the full version in the versioning case, and he can do it at a lower price, these changes in the externalities can even lower his utility. In Section 5 we will examine these welfare issues more closely.

4.3 The free read-only version

When the first constraint is binding, then $v_r = 0$, so the monopoly covers all the market by selling the read-only version at no price (by equation (11), $p_r = [1 - F(v_w)]g(v_r) = 0$, since $g(0) = 0$).

The second constraint must be slack, so $\mu = 0$. Rewriting the first-order conditions (14) and (15) we have that

$$\frac{1 - F(v_w)}{f(v_w)} = v_w, \text{ and} \tag{18}$$

$$\frac{g'(0)}{f(0)} + \frac{\lambda}{[1 - F(v_w)]f(0)} = v_w. \tag{19}$$

The first condition gives a unique solution for v_w, both because of the monotone hazard rate property and of the strict quasi-concavity of the profit function. Let us denote this value by $\widetilde{v}_w$.

First let us prove our by now familiar lemma concerning our sorting condition in v_r.

Lemma 10 *If $0 = v_r < \widetilde{v}_w$, then $g'(0) = g'(v_r) < 1$.*

Proof: Rewriting conditions (18) and (19) and using the fact that $\frac{\lambda}{[1-F(\widetilde{v}_w)]f(0)}$ is positive, we have $\frac{1-F(\widetilde{v}_w)}{f(\widetilde{v}_w)} > \frac{1-F(0)}{f(0)}g'(0)$. Since $0 < \widetilde{v}_w$, by the monotone hazard rate property $\frac{1-F(\widetilde{v}_w)}{f(\widetilde{v}_w)} < \frac{1-F(0)}{f(0)}$, so $g'(0)$ must be less than one. ■

Similarly to the "standard" versioning case, we can see that the monopoly sells less of the full version than if he sells the full version only, and he naturally enlarges the market by $v_r = 0$. Moreover, we can see that this case results in the smallest quantity of the full version defined by the condition (18), which is the solution of the standard single product monopoly problem without network externalities.

Proposition 11 *If $0 = v_r < \widetilde{v}_w$, then $0 = v_r < v_f < \widetilde{v}_w$, and $\widetilde{v}_w$ is bigger than any v_w of the "standard" versioning case.*

Proof: The proof of the first part is by contradiction. Let us suppose that $v_r < \widetilde{v}_w \leq v_f$. Monotone hazard rate property implies $\frac{1-F(v_f)}{f(v_f)} \leq \frac{1-F(\widetilde{v}_w)}{f(\widetilde{v}_w)}$. Replacing each side by using optimum conditions (2) and (18) we have

$$\frac{2}{1 + g'(v_f)}[v_f + g(v_f)] \leq \widetilde{v}_w.$$

But this is a contradiction, since by the concavity of $g(v)$ *and Lemma 10,* $1 + g'(v_f) < 1 + g'(0) < 2$, *and* $v_f + g(v_f) \geq \widetilde{v}_w$.

For the second part suppose that $\widetilde{v}_w \leq v_w$, *for a solution of the "standard" versioning problem. By the monotone hazard rate condition* $\frac{1-F(v_w)}{f(v_w)} \leq \frac{1-F(\widetilde{v}_w)}{f(\widetilde{v}_w)}$. *Replacing the terms by using optimum conditions (16) and (18) yields*

$$v_w + g(v_r) \leq \widetilde{v}_w,$$

which is a contradiction, since in "standard" versioning any $g(v_r) > 0$. ■

We can find a simple explanation for the second result: it can be seen from the price equations (10) and (11) that the price the buyers of the full version are facing is $p_w = [1 - F(v_r)]v_w + p_r$, which is in this case simply $p_w = v_w$, so the monopoly faces his "usual" profitmaximisation problem.

As in the "standard" versioning case, the comparison of the prices is ambiguous. The price of the read-only version ($p_r = 0$) is clearly lower than p_f. However, $p_w = \widetilde{v}_w$ can be larger or less than $p_f = [1 - F(v_f)][v_f + g(v_f)]$.

From equation (19) we have that

$$\lambda = [1 - F(\widetilde{v}_w)][\widetilde{v}_w f(0) - g'(0)] = -\frac{\partial \Pi(\widetilde{v}_w, 0)}{\partial v_r}.$$

For λ to be non-negative, we have to have that $\widetilde{v}_w f(0) - g'(0) \geq 0$. And this necessary condition is sufficient as well, since one of the constraints is slack.

Corollary 12 *The firm introduces the read-only version for free if and only if* $\widetilde{v}_w f(0) \geq g'(0)$, *where* $\widetilde{v}_w$ *is implicitly defined by the equation* $\widetilde{v}_w = \frac{1-F(\widetilde{v}_w)}{f(\widetilde{v}_w)}$.

If this condition is satisfied, it is not profitable for the firm to raise v_r to a positive level. It would decrease the reading base, so writers would not pay such a high price for the full version, and this loss would not be compensated by the small profit made on the readers. From one point of view, this case bears some similarity with the so called shut-down policy of a discriminating monopoly, i.e. choosing to collect profit from high-type consumers only; but on the other hand, the use of the read-only version increases profit to reap from high-types, so it is introduced and sold for free.

4.4 An example

Let us now build a simple example to illustrate the different versioning cases that may occur. Let v to be uniformly distributed on $[0,1]$ so $f(v) = 1$ for all v. Let the valuation for the reading function to be proportional to the valuation for the writing function, $g(v) = kv$. Although we have made the assumption that $g(1) \leq 1$, let us omit this for the sake of the demonstration, so k can take any positive values.[14]

The non-versioning case: We have seen before that this case occurs if and only if $g'(v_f) \geq 1$. Since in this example $g'(v) = k$, that means $k \geq 1$. Using the optimum condition of the full version only case, equation (2), we have that $v_f = \frac{1}{3}$. The equilibrium price can be computed using equation (2): $p_f = \frac{2}{9}(1+k)$.

The free-read only case: We proved that a necessary and sufficient condition for this case to occur is that $\widetilde{v}_w f(0) - g'(0) \geq 0$, where $\widetilde{v}_w$ is implicitly defined by the equation $\widetilde{v}_w = \frac{1-F(\widetilde{v}_w)}{f(\widetilde{v}_w)}$. Solving the latter equation, we have $\widetilde{v}_w = \frac{1}{2}$, and the first condition is satisfied if and only if $k \leq \frac{1}{2}$. By using equation (10), the equilibrium price of the full version is $p_w = \frac{1}{2}$, and by definition p_r and v_r are both 0.

"Standard" versioning: A consequence of the previous points is that this case occurs if and only if $\frac{1}{2} < k < 1$ holds. By solving the system of equations (16) and (17), we have that $v_w(k) = \frac{2}{3}(1-\frac{k}{2})$ and $v_r(k) = \frac{2}{3}(1-\frac{1}{2k})$. equilibrium prices can be derived using equation (10) and (11), $p_w = \frac{1}{9}[(1+\frac{1}{k})(2-k)+(k+1)(2k-1)]$, $p_r = \frac{1}{9}(k+1)(2k-1)$.

By examining the three cases together, we can see the intuition derived in the Kuhn-Tucker analysis to prevail. If the valuation for the reading function is relatively high compared to that for the writing function, i.e. when $k \geq 1$, the firm chooses not to version his product. When the valuation for the reading function gradually decrease, $\frac{1}{2} < k < 1$, the firm continuously increases v_w from $v_w(1) = \frac{1}{3}$, and decreases v_r from $v_r(1) = \frac{1}{3}$. As the valuation for the reading function becomes relatively small to that for the writing function, i.e. when $k \leq \frac{1}{2}$, the firm covers all the market by $v_r(\frac{1}{2}) = 0$, and sets the highest v_w (the smallest quantity of the full version) by $v_w(\frac{1}{2}) = \widetilde{v}_w = \frac{1}{2}$.

[14]Remember that $g'(v)$, which is always k in this example, plays a crucial role in determining which case will occur, so we do not want to restrict this derivative to be less than one. Anyway, we will use this assumption only in examining welfare consequences.

5 Welfare analysis

After analyzing whether the firm chooses to introduce the read-only version, we want to examine how this decision affects the consumers. We have seen from Propositions 9 and 11 that versioning results always in increasing v_w and decreasing v_r, so it creates and destroys some externalities. It may have simultaneously a positive and a negative effect on the welfare of a consumer (only for the one purchasing the full version in both cases), and additionally the price of the product changes as well. By doing the analysis of welfare, we are interested in the question whether there exist some conditions, under which versioning results in a Pareto-improvement, so no one's welfare is decreased.

If versioning occurs, than it was profitable for the firm to choose it, so it is clearly a Pareto-improvement for him. Consumers of type $[0, v_r]$ do not purchase at all in either case (except if $v_r = 0$, but then $U(v_r) = 0$), so their utility is not affected by versioning. Consumers of type $[v_r, v_f]$ did not purchase before, and now they derive a non-negative utility by buying the read-only version , so versioning is a Pareto-improvement for them as well.

Consumers of type $[v_f, v_w]$ would purchase the full version, if it was the only version sold, and they would purchase the read-only version under versioning . The change can be expressed by

$$\Delta U_r(v) = [(1 - F(v_w))g(v) - p_r] - [(1 - F(v_f))(v + g(v)) - p_f].$$

Consumers of type $[v_w, 1]$ would purchase the full version in both case. The change in their utility is the following:

$$\begin{aligned} \Delta U_w(v) &= [(1 - F(v_r))v + (1 - F(v_w))g(v) - p_w] - \\ &\quad -[(1 - F(v_f))(v + g(v)) - p_f]. \end{aligned}$$

Note that since v_w is the indifferent consumer between the full and the read-only version, $\Delta U_w(v_w) = \Delta U_r(v_w)$.

Consider now the following condition:

$$[1 - F(v_w)] + [1 - F(v_r)] > 2[1 - F(v_f)]. \tag{20}$$

The first term of the left-hand side is the amount of full version sold under versioning, so the amount of writing software. The second term is the amount of full version and read-only versions sold under versioning, so the amount of reading software. On the right-hand side we have the amount of full versions software sold, when there is no versioning, which can both write and read. This condition is very similar to the one derived by Schmalensee [1981] as a necessary condition for welfare improvement under price discrimination: the firm's output has to increase. However, in this model the firm's output is not exactly the amount of software he produces (which is always increasing, since $v_r < v_f$), but the amount of the functions the software can perform.

We can see that once this condition is fulfilled, it is sufficient to check the impact of versioning policy on the welfare of the indifferent consumer between the full and the read-only version. If his welfare has been increased, versioning leads to an overall Pareto-improvement.

Proposition 13 *Suppose Condition 20 holds. Then if $\Delta U_w(v_w) = \Delta U_r(v_w) > 0$, versioning leads to a Pareto-improvement.*

Proof: We have to examine only the change of utility of consumers of type $[v_f, 1]$, since in the case of the other agents versioning is clearly a (weak) Pareto-improvement, independently from Condition 20. Consider consumers of type $[v_f, v_w]$. We claim that their gain by versioning, $\Delta U_r(v)$ decreases in v. Indeed,

$$\begin{aligned}\frac{d\Delta U_r(v)}{dv} &= (1-F(v_w))g'(v) - (1-F(v_f))(1+g'(v)) \\ &= -(1-F(v_f)) + (F(v_f) - F(v_w))g'(v) < 0,\end{aligned}$$

since the first term is negative, $F(v_f) - F(v_w)$ is negative and $g'(v) > 0$. Because of that, if $\Delta U_r(v_w) > 0$, then $\Delta U_r(v) > 0$ for all $v \in [v_f, v_w]$.

Consider now consumers of type $[v_w, 1]$. We claim that their gain by versioning, $\Delta U_w(v)$ increases in v. Indeed,

$$\begin{aligned}\frac{d\Delta U_w(v)}{dv} &= 1 - F(v_r) + (1-F(v_w))g'(v) - (1-F(v_f))(1+g'(v)) \\ &= (1-F(v_r)) - (1-F(v_f)) + [(1-F(v_w)) - (1-F(v_f))]g'(v) > \\ &> (1-F(v_w)) + (1-F(v_r)) - 2(1-F(v_f)) > 0,\end{aligned}$$

where the first inequality follows from the fact that $(1-F(v_w)) - (1-F(v_f)) < 0$, $g'(v) < 1$ for all $v \in [v_w, 1]$, since $g'(v_r) < 1$ by Lemmas (8) and (10), and $g(v)$ is concave; the second inequality follows from Condition 20. Because of this, if $\Delta U_w(v_w) > 0$, then $\Delta U_w(v) > 0$ for all $v \in [v_w, 1]$. ■

An intuitive sufficient condition for the positivity of $\Delta U_w(v_w)$ would be the price fall of the full version, since condition 20 suggests a welfare improvement from the externality restructuring at a smaller price.[15] This conjecture is confirmed by the following lemma.

Lemma 14 *If $p_w < p_f$ and condition 20 holds, then $\Delta U_w(v_w) > 0$.*

Proof: Expanding $\Delta U_w(v_w)$, we have that

$$\begin{aligned}\Delta U_w(v_w) &= [(1-F(v_r))v_w + (1-F(v_w))g(v_w) - p_w] - \\ &\quad -[(1-F(v_f))(v_w + g(v_w)) - p_f] \\ &> [(1-F(v_r)) - (1-F(v_f))]v_w - \\ &\quad -[(1-F(v_f)) - (1-F(v_w)]g(v_w) \\ &> (1-F(v_w)) + (1-F(v_r)) - 2(1-F(v_f)) > 0,\end{aligned}$$

where the first inequality follows from $p_w < p_f$, the second from $g(v_w) < 1$ and $(1-F(v_f)) - (1-F(v_w) > 0$ and the third from Condition 20. ■

To summarize, we can state the following sufficient condition for Pareto-improvement:

Corollary 15 *If the amount of the functions the software can perform increases and price of the full version falls due to versioning, then it leads to a Pareto-improvement.*

6 Conclusion

We have examined a software monopoly's special differentiation policy, namely introducing a functionally degraded version of his "top product" by removing the writing function. We have derived the necessary and sufficient conditions for the profitable supply of a read-only version along with the full version, and the conditions of providing it for free. We have examined the welfare implications of this type of software versioning, but we have not found unambiguous results. However, some conditions have been derived when the introduction of the read-only version will lead to a strict Pareto-improvement.

[15]These two conditions together can be quite restricitive with some specifications, they are never fulfilled for example with the linear model studied in subsection 4.4.

Although the continuous type model we used is quite general, let us mention some caveats to the analysis. First, to be able to handle the coordination problem consumers are facing, we have assumed perfect and positive correlation between the consumers' valuations for the different software functions. Working with more general two-dimensional distributions may give more general results, but it would involve a much more complicated analysis, mainly because the problem, which arises from the multiplicity of equilibria cannot be so easily solved.[16] Second, the software were assumed to be pure network goods, i.e. consumers have no stand-alone valuations for using it. This property may fit text word processing software, like Acrobat or Scientific Workplace, but one may encounter other type of software, like Mathematica or RealPlayer, which can have functions that are useful outside of a network as well. Including the stand-alone values would add one more dimension to the space of characteristics. Third, the firm may differentiate his product even further, since a software is usually a bundle of more than just two functions, enlarging again the characteristic space. Examples of this kind can also be found, Scientific Word, for example, misses some of the mathematical tools Scientific Workplace is using, but one may still write mathematical documents by using it.

Software are by nature durable goods, but this property was completely excluded in the analysis.[17] On the top of the mathematical difficulties that may arise by including dynamics in the model, there are some other problems, which have to be considered. First, there are two products the user can buy, so consumers may switch in later periods not only from not buying to buying, but also between different versions (naturally only from read-only to full version). Second, software firms continuously upgrade their products, sometimes even in an excessive manner[18]; the recent Acrobat and Scientific Workplace are both version 5.0, for example. Upgrades raise the question of compatibility as well: usually, the newer versions are only backward compatible, which degrades the value of the old version, forcing consumers to buy the new version.[19] There is another aspect

[16] A user's guide for multidimensional screening was given by Armstrong and Rochet [1999], more detailed analysis is given by Armstrong [1996] and Chone and Rochet [1998].

[17] For durable goods exhibiting network externalities and the Coase-conjecture in this framework see, for example, Cabral et al. [1999] and Economides [2000].

[18] On this issue, see Ellison and Fudenberg [2000].

[19] This type of planned obsolescence of a software was first modeled by Choi [1994].

of upgrades: the software firms may find it profitable to offer upgrade discounts, by which they can extract some information about users.[20]

Last, the monopoly situation was not questioned. The firm may face the threat of potential entrants into the market, and if they do, some form of competition takes place. Most of the network industries involve large sunk costs, so entry is quite costly. The incumbent firm has his installed base advantage, and he may sacrifice some of his monopoly profit to increase this installed base, deterring the potential entrants from entry.[21] Additionally, the question of compatibility arises again, now on both sides of the market: some examples show that entrants accept the standard developed by the incumbent (the PDF format of Acrobat, for example), so they have a large reading base, and try to undercut the price of the incumbent's full version.[22] Including these extensions are left for future research.

[20] Fudenberg and Tirole [1998] analyse this possibility.

[21] Fudenberg and Tirole [2000] have built a model to explain this phenomenon.

[22] Jullien [2001] examines a competitive game between a dominant and a challenging network in the presence of network externalities. His model, although in an other context, deals with a lot of questions addressed in this paper, namely (price) discrimination, cross-subsidization and compatibility.

References

[1] Adams, W., and J. Yellen [1976]: "Commodity Bundling and the Burden of Monopoly", *Quarterly Journal of Economics* 90, 475-498.

[2] Armstrong, M. [1996]: "Multidimensional Nonlinear Pricing", *Econometrica* 64, 51-75.

[3] Armstrong, M. and J-C. Rochet [1999]: "Multidimensional Screening: A User's Guide", *European Economic Review* 43, 959-979.

[4] Bakos, Y., and Brynjofsson, E. [1999]: "Bundling Information Goods, Pricing, Profits, and Efficiency", *Management Science* 45, December.

[5] Bakos, Y., and Brynjofsson, E. [2000]: "Bundling and Competition on the Internet", *Marketing Science* 19, Winter.

[6] Cabral L., D. Salant, and G. Woroch [1999] "Monopoly Pricing with Network Externalities, *International Journal of Industrial Organization* 14*, 837-855.*

[7] Choi, J. [1994], "Network Externality, Compatibility Choice, and Planned Obsolescence", *Journal of Industrial Economics* 17, 167-181.

[8] Chone, P., and J-C. Rochet [1998]: "Ironing, Sweeping, and Multidimensional Screening", *Econometrica* 66, 783-826.

[9] Deneckere, R., and P. McAfee [1996]: "Damaged Goods", *Journal of Economics and Management Strategy* 5, 149-174.

[10] Economides, N. [2000]: "Durable Goods Monopoly with Network Externalities with Application to the PC Operating System Market", *Quarterly Journal of Electronic Commerce* 1.

[11] Ellison, G., and D. Fudenberg [2000]: "The Neo-Luddite's Lament: Too Many Upgrades in the Software Industry", *Rand Journal of Economics* 31.

[12] Farrel, J., and G. Saloner [1985]: "Standardization, Compatibility, and Innovation", *Rand Journal and Economics* 16, 70-83.

[13] Fudenberg, D., and J. Tirole [1998]: "Upgrades, Tradeins and Buybacks", *Rand Journal of Economics* 29, 235-258.

[14] Fudenberg, D., and J. Tirole [2000]: "Pricing a Network Good to Deter Entry", *Journal of Industrial Economics* 23, 373-390.

[15] Hahn, J. [2001]: "Functional Quality Degradation of Software with Network Externalities", mimeo Keele University.

[16] Jing, B. [2000]: "Versioning Information Goods with Network Externalities", mimeo University of Rochester.

[17] Jullien, B. [2001]: "Competing with Network Externalities and Price Discrimination", mimeo IDEI, University of Toulouse.

[18] Katz, M., and C. Shapiro [1985]: "Network Externalities, Competition, and Compatibility", *American Economic Review* 75, 424-440.

[19] Maskin, E., and J. Riley [1984]: "Monopoly with Incomplete Information", *Rand Journal of Economics* 15, 171-196.

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

[21] Schmalensee, R. [1981]: "Output and Welfare Implications of Monopolistic Third-Degree Price Discrimination", *American Economic Review* 71, 242-247.

[22] Shapiro, C., and H. Varian [1999]: *Information Rules*, Harvard Business School Press.

Appendix

Proof of Lemma 1

We have seen that for all price p_f the firm sets, there exists only one v_f that forms an equilibrium, let us denote this type by $h(p_f)$. The monopoly's profit is $\Pi(p_f) = p_f[1 - F(h(p_f))]$, so maximization in p_f gives first-order condition $1 - F(h(p_f)) - p_f f(h(p_f)))h'(p_f) = 0$.

Now consider the case when the monopoly's profit is maximized as a function of v_f. First note that the monopoly would never choose a v'_f, for which there exists no p_f such that $v'_f = h(p_f)$, i.e. which is not on the demand function. Indeed, since at the price p'_f satisfying equation (1) there exists another $v''_f < v'_f$, which gives higher utility to every consumer, so v'_f cannot be an equilibrium by the Pareto criterion.

The inverse demand function can thus be expressed: $p_f = h^{-1}(v_f)$. The firm's profit in v_f is $\Pi(v_f) = h^{-1}(v_f)[1-F(v_f)]$. Maximization in v_f gives the first-order condition $(h^{-1}(v_f))'[1 - F(v_f)] - h^{-1}(v_f)f(v_f) = 0$, which is equivalent with the first-order condition of the former problem, since $(h^{-1}(v_f))' = \frac{1}{h'(p_f)}$. The two ways of maximization give thus the same solution. *Q.E.D.*

Proof of Lemma 2

Consider first consumers with a type $v > v_w$. Substituting equation (7) in condition (3) and equation (9) in condition (5) give

$$\begin{aligned} (n_f^e + n_r^e)(v - v_w) + n_f^e(g(v) - g(v_w)) &\geq 0, \text{ and} \\ (n_f^e + n_r^e)(v - v_w) &\geq 0, \end{aligned}$$

which are clearly satisfied by the monotonicity of $g(\cdot)$, so they all purchase the full version.

Consider now consumers with a type $v_r < v < v_w$. Substituting equation (7) in condition (5) and equation (8) in condition (6) give

$$\begin{aligned} (n_f^e + n_r^e)(v_w - v) &\geq 0, \text{ and} \\ n_f^e(g(v) - g(v_r)) &\geq 0, \end{aligned}$$

which are both satisfied, so they all purchase the read-only version.

Last, consider consumers with a type $v < v_r$. By purchasing the full version they would derive a utility of $(n_f^e + n_r^e)(v - v_w) + n_f^e(g(v) - g(v_w))$, or by buying the read-only version their utility would be $n_f^e(g(v) - g(v_r))$, which are both negative, so they do not purchase at all. *Q.E.D.*

Proof of Lemma 4

By following the same steps as in Lemma 1, it can be seen that the first-order conditions of the two problem are equivalent where the demand function $h(p_w, p_r)$ is invertible.

However, since the function $h(p_w, p_r)$ does not take all the values in $[0,1] \times [0,1]$, it may be possible that the problem written in marginal types gives a solution (v'_w, v'_r), for which there exists no (p_w, p_r), such that $h(p_w, p_r) = (v'_w, v'_r)$. But this pair would not form an equilibrium, since at the prices (p'_w, p'_r) that satisfy (10) and (11) at (v'_w, v'_r), there exists another pair of larger quantities, $(v''_w, v''_r) < (v'_w, v'_r)$, such that $(v''_w, v''_r) = h(p'_w, p'_r)$, and it gives higher utility of every consumer, so (v'_w, v'_r) would be chosen. Hence the two maximization problems give the same solutions. *Q.E.D.*

Proof of Lemma 5

We will prove the strict quasi-concavity of the function in all the pairs $(v_w, v_r) \in [0,1) \times [0,1)$, where $\nabla\Pi(v_w, v_r) \neq 0$. This equality is satisfied only in the so called "standard" versioning optimum, and for this case we will check the second-order condition in the next proof. Since it will be satisfied, the function is strictly concave in that point, which implies strict quasi-concavity.

The function $\Pi(v_w, v_r) = [1 - F(v_w)][1 - F(v_r)][v_w + g(v_r)]$ is strictly quasi-concave if the Hessian matrix $D^2\Pi(v_w, v_r)$ is negative definite in the subspace $\{z \in R^2 : \nabla\Pi(v_w, v_r) \cdot z = 0\}$ for all (v_w, v_r). This is true if and only if the bordered Hessian has a positive determinant, i.e.

$$D = 2\frac{\partial\Pi(v_w, v_r)}{\partial v_w}\frac{\partial\Pi(v_w, v_r)}{\partial v_r}\frac{\partial^2\Pi(v_w, v_r)}{\partial v_w \partial v_r} - [\frac{\partial\Pi(v_w, v_r)}{\partial v_w}]^2\frac{\partial^2\Pi(v_w, v_r)}{\partial v_r^2}$$

$$-[\frac{\partial\Pi(v_w, v_r)}{\partial v_r}]^2\frac{\partial^2\Pi(v_w, v_r)}{\partial v_w^2} > 0.$$

Let us introduce two new variables, D_{v_w} and D_{v_r} to simplify computations:

$$\begin{aligned}
\frac{\partial \Pi(v_w, v_r)}{\partial v_w} &= [1 - F(v_r)][1 - F(v_w) - f(v_w)(v_w + g(v_r))] = [1 - F(v_r)]D_{v_w}, \\
\frac{\partial \Pi(v_w, v_r)}{\partial v_r} &= [1 - F(v_w)][(1 - F(v_r))g'(v_r) - f(v_r)(v_w + g(v_r))] = \\
&= [1 - F(v_w)]D_{v_r}.
\end{aligned}$$

Note that $\nabla \Pi(v_w, v_r) = 0$ only if $D_{v_w} = D_{v_r} = 0$, since the other term is always positive.

By using these new variables, the other derivatives are the following:

$$\begin{aligned}
\frac{\partial^2 \Pi(v_w, v_r)}{\partial v_w \partial v_r} &= -f(v_r)[1 - F(v_w)] - f(v_w)D_{v_r} = \\
&= -f(v_w)[1 - F(v_r)]g'(v_r) - f(v_r)D_{v_w}, \\
\frac{\partial^2 \Pi(v_w, v_r)}{\partial v_w^2} &= [1 - F(v_r)][-2f(v_w) - f'(v_w)(v_w + g(v_r))], \\
\frac{\partial^2 \Pi(v_w, v_r)}{\partial v_r^2} &= [1 - F(v_w)][(1 - F(v_r))g''(v_r) - 2f(v_r)g'(v_r) - \\
&\quad -f'(v_r)(v_w + g(v_r))].
\end{aligned}$$

Expanding D we have the following expression:

$$\begin{aligned}
D &= [1 - F(v_w)][1 - F(v_r)]D_{v_w}D_{v_r}[-f(v_r)D_{v_w} - f(v_w)D_{v_r} - \\
&\quad -2f(v_w)f(v_r)(v_w + g(v_r))] - \\
&\quad -[1 - F(v_w)]^2[1 - F(v_r)]D_{v_r}^2[-2f(v_w) - \underbrace{f'(v_w)(v_w + g(v_r))}] - \\
&\quad -[1 - F(v_r)]^2[1 - F(v_w)]D_{v_w}^2[\underbrace{(1 - F(v_r))g''(v_r)} - \\
&\quad -2f(v_r)g'(v_r) - \underbrace{f'(v_r)(v_w + g(v_r))}].
\end{aligned}$$

We can simplify by $2[1-F(v_w)][1-F(v_r)]$, and since all the terms are positive, it does not change the sign of the expression. Let us omit the highlighted terms (with their respective signs), so the expression decreases, since they are all positive according to the assumptions made before ($f\prime(v) > 0, g''(v) < 0$). Combining the terms we have that

$$\begin{aligned}
D' &= -D_{v_w}D_{v_r}f(v_w)f(v_r)(v_w + g(v_r)) + D_{v_r}^2 f(v_w)[(1 - F(v_w)) - \frac{D_{v_w}}{2})] + \\
&\quad +D_{v_w}^2 f(v_r)[(1 - F(v_r))g'(v_r) - \frac{D_{v_r}}{2})] \geq
\end{aligned}$$

$$
\begin{aligned}
&\geq -D_{v_w}D_{v_r}f(v_w)f(v_r)(v_w+g(v_r)) + D_{v_r}^2[f(v_w)]^2(v_w+g(v_r)) + \\
&\quad +D_{v_w}^2[f(v_r)]^2(v_w+g(v_r)) \geq \\
&\geq (v_w+g(v_r))[D_{v_w}f(v_r) - D_{v_r}f(v_w)]^2 \geq 0.
\end{aligned}
$$

Since $D > D'$, D is positive. *Q.E.D.*

Second order conditions for the "standard" versioning case

We have to check that under first-order conditions (16) and (17) the Hessian of the objective function $\Pi(v_w, v_r)$ is positive definite.

$$\frac{\partial^2\Pi(v_w, v_r)}{\partial v_w^2} = [1 - F(v_r)][-2f(v_w) - f'(v_w)(v_w + g(v_r))] < 0$$

is always satisfied, we have to check if the determinant of the Hessian is positive:

$$D = \frac{\partial^2\Pi(v_w, v_r)}{\partial v_w^2}\frac{\partial^2\Pi(v_w, v_r)}{\partial v_r^2} - [\frac{\partial^2\Pi(v_w, v_r)}{\partial v_w \partial v_r}]^2$$

Using the results of the previous proof, and the fact that in this case $D_{v_w} = D_{v_r} = 0$, we have that

$$
\begin{aligned}
[\frac{\partial^2\Pi(v_w, v_r)}{\partial v_w \partial v_r}]^2 &= [1 - F(v_w)][1 - F(v_r)]f(v_w)f(v_r)g'(v_r), \\
\frac{\partial^2\Pi(v_w, v_r)}{\partial v_w^2}\frac{\partial^2\Pi(v_w, v_r)}{\partial v_r^2} &= [1 - F(v_w)][1 - F(v_r)] * [-2f(v_w) - \\
&\quad - \underbrace{f'(v_w)(v_w + g(v_r))}][\underbrace{(1 - F(v_r))g''(v_r)} - \\
&\quad -2f(v_r)g'(v_r) - \underbrace{f'(v_r)(v_w + g(v_r))}] \\
&> 4[1 - F(v_w)][1 - F(v_r)]f(v_w)f(v_r)g'(v_r),
\end{aligned}
$$

since the highlighted terms are all negative with their signs by the assumptions. So $D > 0$. *Q.E.D.*