

# Product development and pricing strategy for information goods under heterogeneous outside opportunities\*

Ying-Ju Chen and Sridhar Seshadri<sup>†</sup>

Stern School of Business, New York University

## Abstract

This paper considers a two-stage development problem for information goods with costless quality degradation. In our model, a seller of information goods faces customers that are heterogeneous with regard to both the marginal willingness to pay for quality and the outside opportunity. In the development stage, the seller determines the quality limit of the product. In the second stage, the seller's problem is to design the price schedule corresponding to different quality levels, taking into account production and distribution costs.

We show that versioning is optimal for the seller when customers have multiple outside options, or more generally, convex reservation utilities. In addition, we show that in the optimal solution, the seller discards both low-end and high-end customers. Among those that are served, the seller offers a continuum of (inferior) versions to customers with relatively low willingness to pay, and extracts full information rent from each of them. A common version with the quality limit is offered to the rest.

We further prove that the seller should offer a single version when reservation utilities are either concave or linear. Through numerical experiments, we study the sensitivity of our results to changes in the cost structure and customer utilities.

**Keywords:** versioning, quality degradation, price discrimination, information goods, heterogeneous outside opportunities

## 1 Introduction

The value of a digital good is measured by its information content rather than its physical content. Hence, it is usually hard to produce the first copy, but easy to reproduce and distribute the product. These characteristics of “information goods” imply that they have a fairly specific cost structure: high fixed cost for product development but zero or near zero marginal cost for production and distribution. For example, the cost of establishing a database such as the Wharton Research Data Services (WRDS) is considerably high, but

---

\*We thank Anindya Ghose, Ke-Wei Huang, and Arun Sundararajan for many stimulating discussions.

<sup>†</sup>Address for correspondence: 44 W4 St, NY, NY 10012; e-mail:{ychen0,sseshadr}@stern.nyu.edu.

once it is established, the marginal cost of serving additional subscribers is not very high. Music recording, movies, internet search engines, on-line content providers, and journals all share the same characteristics, since sellers do not have to incur significantly more cost for allowing one more subscription or download.

Similarly, software is another product that represents and justifies the nomenclature of information good. For example, to design a scientific application software such as Mathematica, one has to consider the ability to import/export data, the flexibility of working on multiple platforms (Linux, Mac, and Windows), and the ability to carry out symbolic operations, produce graphic outputs, and do sophisticated numerical investigations. The decisions to be made in developing an antivirus software such as Symantec Norton include multi-layer delegation, timely virus code updating, Firewall/VPN function designs, email verification and attack prevention, and website filtering as well as antispam. Numerous features are carefully considered for inclusion in the development stage for scientific software such as ILOG, Maple, Matlab, Minitab, and antivirus software (Kaspersky and TrendMicro), and other functional software packages. Compared to the cost of developing the above features, the production and distribution costs are negligible.<sup>1</sup>

The second and perhaps more subtle characteristic of information goods is that degradation, i.e., reduction, of the product's quality is relatively easy and inexpensive. For example, the degradation of software typically involves disabling a subset of functions, inserting incompatibility with respect to accessing contemporary software, closing the access to high level databases, introducing intentional delay, and providing restricted technical support. Examples of software that utilize some of these degradation techniques include Adobe Acrobat, ILOG, Norton Antivirus, Visio, and numerous others. These degradations can be done by minor manipulation of the features without any physical change to the product. Therefore, the cost is negligible compared to the amortized development and marketing related costs. Internet service also has the same characteristic. Many Internet services offer free-sponsored sites that provide daily news as well as fee-based sites that convey more specific or detailed information, for example, AOL, Classmates.com, CNN.com, and Yahoo (Riggins

---

<sup>1</sup>For example, it typically costs less than \$1.50 to make one more CD copy, see Raghunathan [2000].

[2002]). Versioning may also be implemented by the choice of timing, (e.g., InterQuote and PAWWS), or by introducing banner advertisement to discomfort users (Eudora, Kazaa, and Silicon investor membership).

To capitalize on these two specific characteristics of information goods, information providers usually develop a high-end (flagship) product in the “development” phase. After the technological quality limit has been established with the development of a flagship product, in the “production/distribution” phase the sellers degrade the product to provide quality-differentiated versions. This is known as the “versioning” strategy. This strategy, which Shapiro and Varian [1998] call “the smart way to sell information”, is believed to be profitable in both software and information services industries. In this paper, we use the word “quality” in the following way: Like other goods, the quality of software and information goods is multi-dimensional. For example, these dimensions can be delay, user interface, convenience, image resolution, speed of operation, format, capability, features, comprehensiveness, annoyance, and support (Shapiro and Varian [1998]). Juran and Gryna [1951] propose a methodology to aggregate quality measures on multiple dimensions such as these into a single quality score. Using this approach, each user aggregates the measure of quality on different dimensions into a score using weights that reflect the relative importance of the dimensions. It is usual to assume that the weights are the same for all customers within a target segment. Thus, the quality limit can be viewed as the maximum quality score that a customer can obtain from the combination of features of the product.

In this paper, we investigate the design and distribution problem of information good providers that serve individual users. Unlike institutional users who may purchase multiple licenses or subscriptions of the same product, an individual user typically demands at most one unit of the product. The key ingredients of the design and distribution problem are: (1) The cost of developing the product, which might depend on the quality limit, i.e., the set of features, functionality, etc. (2) The production, distribution, and servicing cost (which may or may not depend on quality). (3) The utility derived by customers. (4) Customers’ reservation values. We discussed the high fixed cost of development above, and hence in the sequel describe the other elements. Following this, we discuss the prior research in this area.

In the production stage, since the quality degradation is relatively easy (as for the information goods such as software and information service), the variable cost does not increase significantly in the quality of the product. Thus, it is reasonable to assume that the variable cost is independent of the quality level. Therefore, the marginal production cost is usually neglected in the analysis. Other than the cost of producing the copy, the seller might incur a distribution cost for each sale transaction. This expenditure might comprise the cost of obtaining a new subscription, the handling fee for accepting returns/exchanges, and other administrative/marketing-related costs. In all these activities, costs can usually be allocated on a per customer basis. The literature on versioning of information goods usually ignores these variable costs for simplicity.

Customers' utilities are an important input to the versioning strategy. Since versions are quality-differentiated (high-end products may contain more functions, allow switching among platforms, be compatible to more products), it is normal to assume that all customers unambiguously prefer higher quality to lower quality. Nevertheless, customers differ in unobservable preferences (types) with regard to quality. This inherent heterogeneity among customers induces the seller to offer different versions at different prices. Customers self-select after evaluating the cost-benefit trade-off amongst these versions. Following the celebrated monopoly pricing paper by Mussa and Rosen [1978], the majority of papers on pricing information goods model heterogeneity by allowing customers to have different but constant marginal willingness to pay for quality, see Bhargava and Choudhary [2001, 2004], Jing [2006], and Jones and Mendelson [1998]. Ghose and Sundararajan [2005] and Raghunathan [2000], who adopt nonlinear (quadratic) utility functions, are exceptions to this literature.

The last ingredient is the heterogeneity in the customers' outside opportunity, which gives rise to type-dependent reservation utilities. The heterogeneity of reservation utilities could arise due to several reasons. Three specific reasons are important for the information good versioning problem: (1) The option to use off-the-shelf substitute products from competing firms. (2) Options created by the recourse to self-developed solutions. (3) The option to use a custom designed product. As an example, suppose a user needs to perform numerical investigations. She might be leaning towards purchase of Mathematica from Wolfram

Research Inc. But, she is also aware that she can alternatively purchase scientific software from Scientific Workplace or Symbolic Math Toolbox. She could also write a C++/Matlab program or continue to use the customized product (e.g., Maple) that she already subscribes to. As another example, investors that wish to subscribe to PAWWS will consider subscribing to Bloomberg, Reuters, or RiskView, or evaluate searching financial information online as required, and self-customization using MSN Money or Yahoo! Finance. These reasons lead to convex reservation utilities, and will be considered in the main model (Secs. 2-5).

Outside opportunities also can be type-dependent due to the differences in the cost of adapting to or switching to a new product. Different customers may have different existing subscription levels, technical sophistication, and software/hardware. Due to these differences, customers might experience different switching costs. Switching costs could also be related to compatibility with the products currently used by the customers. However, switching cost need not always result in convex reservation utilities. The concave and linear cases are discussed separately in Section 6.

Thus, “heterogeneous outside opportunities” in our view is probably more the norm than the exception for information goods. Other writers have observed this phenomenon for information goods. For example, Sundararajan [2004] introduces outside opportunity as the chance of obtaining a pirated version of the software, and Huang and Sundararajan [2005] interpret the reservation utility as the effort required to self-develop the product.

Many papers have investigated the profitability of versioning information goods using some (but not all) of the above ingredients. Most papers focus on the second-stage problem of versioning given a finite set of quality levels (already developed in the first stage), and in fact argue that the variable cost is either zero or concave in quality. They typically adopt constant marginal willingness to pay to model customer preferences and do not consider the heterogeneity of outside opportunities. With these assumptions, Bhargava and Choudhary [2001] find the optimal solution is to create a single version, see also Jones and Mendelson [1998] and Weber [2002]. Bhargava and Choudhary [2001] suggest in the conclusion to their paper that additional factors must be included to justify the versioning strategy. Our model incorporates heterogeneous reservation utilities, and demonstrates that it is an important

Table 1: Summary of previous literature on versioning information goods

| Paper                         | Utility   | Variable cost | Other features  | No. of versions |
|-------------------------------|-----------|---------------|-----------------|-----------------|
| Bhargava and Choudhary [2001] | linear    | concave       |                 | one             |
| Bhargava and Choudhary [2004] | linear    | 0             | externality     | two             |
| Ghose and Sundararajan [2005] | quadratic | 0             | empirical       | multiple        |
| Jing [2006]                   | linear    | 0             | externality     | two             |
| Jones and Mendelson [1998]    | linear    | 0             | competition     | one             |
| Raghunathan [2000]            | quadratic | 0             | sequential game | multiple        |

factor for versioning of information goods.

Versioning has been reported as a profitable strategy when the information goods convey network effects, see Bhargava and Choudhary [2004] and Jing [2006]. These papers predict that the seller should provide exactly two versions. Deviating from the standard utility setting, some authors assume nonlinear (quadratic) utility functions, which induces versioning (Ghose and Sundararajan [2005] and Raghunathan [2000]). Table 1 summarizes recent findings regarding versioning, along with assumptions regarding utilities, costs, etc.

Both Bhargava and Choudhary [2004] and Jing [2006] argue that it is optimal to offer exactly two versions when there are network effects. Therefore, network effects cannot explain the multiplicity of versions of products, such as MS Encarta, MS Windows XP, Quickbooks, TurboTax, and PC-Cillin (Ghose and Sundararajan [2005]). Multiplicity of versions is widely observed in the information services industry as well (e.g., Reuters.com provides different packages for financial market professionals, corporate customers, and media professionals). When customers' utility function is nonlinear, the number of versions could depend on the number of distinct customers' segments (types). This result has been known in the product design literature, see, e.g., Maskin and Riley [1984]. However, the added effect of nonlinearity on versioning has not been previously studied. We do this by solving the two-stage problem of development and production for both the standard and nonlinear utility function cases and compare the solutions. We show via numerical experiments that the structure of the optimal solution to the versioning problem carries over to the nonlinear case for a wide range of problem parameters.

In this paper, we first justify the rationale for creating multiple versions of information goods under the standard utility setting and without network effects. We do not deny that these other effects exist, and that they are important. However, we believe our approach provides useful design principles that are different from the ones identified in the earlier work due to the insight provided by joint solution to the two-stage problem. Since we model all ingredients described above, our versioning solution also differs quite a bit from the previous solutions, as described next.

We show that versioning could be profitable because customers have multiple outside options, or more generally, convex reservation utilities, even though they have constant marginal willingness to pay and network effects do not contribute to the utilities. We show that the convexity of reservation utility is a direct consequence of the multiplicity of outside options. Hence, a slight perturbation of standard setting induces the sellers to adopt versioning. The implication for providers of information goods is that careful investigation of available outside options facing their target customers is critical before implementing versioning. For example, the inherent difference of customers' outside opportunities could potentially explain why some software providers sell a single quality to all customers (Microsoft Visual FoxPro, Nova PhotoImpact, and UBI Soft's Red Steel) but others offer multiple versions (Mathematica, Norton Antivirus, and TurboTax), even though they have similar cost structures for product development and network effects are present.<sup>2</sup> When there is no or only one effective outside option accessible to target customers, a single version is preferred by the seller; if multiple options co-exist, the seller can increase her revenue by offering multiple versions.

We also show that, in addition to offering multiple versions, the seller will discard both the low-end and the high-end customers. This peculiar exclusion result is new to the information goods and more generally, the product design literature. Among those served, the seller extracts full information rent from customers with relatively low willingness to pay, and offers a common version (the quality limit) to the rest. Note that this full rent

---

<sup>2</sup>The information regarding the number of versions of all examples cited in this paper was updated last on January 14, 2007. These include Microsoft Visual FoxPro (<http://www.amazon.com/software>), Nova PhotoImpact (<http://www.novadevelopment.com/>), UBI Soft's Red Steel (<http://store.ubi.com/>), Mathematica (<http://www.wolfram.com/products/>), Norton Antivirus (<http://www.symantec.com/product/index.jsp>), and TurboTax (<http://turbotax.intuit.com/>).

extraction of a continuum of low-end customers is not reported as an equilibrium outcome when versioning results from either network effects or nonlinearity of utilities. The result that a continuum of high-end customers obtain the product with quality limit is also in contrast to the predictions offered by introducing network effects or nonlinear utilities: in both cases, efficiency is achieved only for the highest-type customer.

We provide a simple rule for selecting the optimal quality limits to achieve either first-degree or second-degree discrimination. When the quality limit is pre-determined, the information asymmetry forces the seller to give up some transactions that are efficient in the first-best scenario. As the quality limit is raised, the seller gathers a strictly higher profit under both forms of price discrimination. The existing literature often ignores the distribution cost for simplicity. Our analysis does not rule it out and hence is more general. We demonstrate that the profitability of versioning is independent of this effect, and the simplification of assuming no distribution cost can be incorporated as a special case. The exclusion point of low-end customers is completely determined by the reservation utility and distribution cost, independent of other inputs such as the quality limit and the distribution of customers' types. Nevertheless, the starting point of offering the flagship product is jointly determined by all above primitives. This demonstrates the different degrees of sophistication needed in order to identify these two key thresholds, which might serve as handy guidelines for practitioners in the information goods industry.

We then extend our studies to incorporate linear or concave variable costs and concave reservation utilities. We show that if the reservation utility is linear or concave, versioning is suboptimal. We then numerically examine the optimal quality schedule with convex reservation utilities but concave variable costs. In these experiments, we examine the sensitivity of our results with regard to variable costs, customers' utilities, and their reservation values. The managerial implications of these findings are discussed in Section 6.

The rest of this paper is organized as follows. In Section 2, we introduce the model. Section 3 considers the scenario where the seller is able to observe the customers' willingness-to-pay, and in Section 4 this becomes customers' private information, and hence the seller has to offer a menu to induce self-selection. In Section 5, we discuss some comparative statics



for both informational scenarios. Section 6 provides some simulation results and managerial insights, and Section 7 concludes.

## 2 Model

In our model, a seller of information goods faces customers that possess heterogeneous willingness to pay on the quality. The product development takes place sequentially in two stages: the development stage and the production/distribution stage. In the development stage, the seller chooses the quality limit  $\bar{q}$  by devoting a deterministic convex cost  $C(\bar{q})$ . This formulation captures the various decisions she might have to make to impact the quality limit. We assume that degradation is costless, and hence the seller can provide any quality level  $q \in [0, \bar{q}]$  in the production stage, without incurring any extra cost of reengineering/redeveloping. There is a distribution cost  $c(q) \equiv c$  if a product is sold to a customer, independent of the product's quality. That is, if the seller sells a product with quality  $q$  to a customer, her net payoff will be  $\pi(q) = p(q) - c$ ,  $\forall q \in [0, \bar{q}]$ , where  $p(q)$  is the money transfer between the seller and the customer. The seller's problem is to first find an optimal target quality level  $\bar{q}$ , and then propose a menu of quality/price bundles to these customers.

Customers' willingness to pay is assumed to be of the linear, separable format  $u(q, \theta) = \theta q - p(q)$ , where  $\theta$  is the user's marginal willingness to pay (type) with distribution function  $F(\theta)$  and its density  $f(\theta)$  over a finite support  $[0, R]$ . The value of  $R$  captures the maximum marginal willingness to pay for quality and the extent of market heterogeneity of customer preferences on quality. The seller knows the utility function  $u(q, \theta)$ , the entire distribution  $F(\theta)$ , but she is unable to observe customers' types.

We assume that multiple outside options are available if customers refuse to purchase from the seller. Each outside option is characterized by the quality level  $q$  and its associated nonnegative cost  $s(q)$  with  $s(0)$  normalized to 0. These outside options may refer to the substitute products offered by competing firms, the self-developed solutions, or the customized products customers have already been endowed with. In the above scenarios, the cost  $s(q)$  corresponds to respectively the price paid for purchasing a competing product, the self-development cost, and the subscription fee paid for the customized product.

The reservation utility  $r(\theta)$  corresponds to a customer's payoff after self-selecting her

favorite outside option, which has the desired properties as shown below. If all customers unambiguously prefer one option, then we say there exists only one effective option.

**Lemma 1.** *Suppose customers are free to select from a number of outside options  $\{q, s(q)\}$ , and their reservation utility follow from this alternative. Then  $r(\theta)$  is increasingly convex in  $\theta$ , and  $r(0) = 0$ .  $r(\theta)$  is linear if and only if there exists only one effective option.*

*Proof.* By definition  $r(\theta) = \max_{q \geq 0} \{\theta q - s(q)\}$ . Consider two types  $\theta_1$  and  $\theta_2$ . Assume that the maximizer for type- $\theta_1$  customer is  $q_1 \geq 0$ , i.e.,  $r(\theta_1) = \theta_1 q_1 - s(q_1)$ . We have  $r(\theta_2) = \max_{q \geq 0} \{\theta_2 q - s(q)\} \geq \theta_2 q_1 - s(q_1) = \theta_1 q_1 - s(q_1) + q_1(\theta_2 - \theta_1) = r(\theta_1) + q_1(\theta_2 - \theta_1)$ , and therefore  $r(\theta_2) \geq r(\theta_1) + q_1(\theta_2 - \theta_1), \forall \theta_1, \theta_2$ . Note that the above inequality holds for arbitrary pair of types  $\theta_1, \theta_2$ , and hence it implies that  $r(\theta)$  is convex in type.

We now show that  $r(\theta)$  is monotonic. Without loss of generality we assume  $\theta_1 \leq \theta_2$ . The optimality condition leads to  $r(\theta_2) \geq \theta_2 q_1 - s(q_1) = \theta_1 q_1 - s(q_1) + q_1(\theta_2 - \theta_1) \geq r(\theta_1)$ , where the last inequality follows from  $q_1$  being nonnegative. Therefore  $r(\theta)$  is increasing.

The only possibility for  $r(\theta)$  being linear occurs when the maximizer  $q$  is identical for all types, in which case a unique outside option dominates all other options unambiguously. This case degenerates to the single option scenario.  $\square$

This lemma implies that with multiple non-degenerate outside options, the reservation utility  $r(\theta)$  will inevitably be convex.<sup>3</sup> We believe “multiple outside options”, thus “convex reservation utility”, is the natural model setting. We therefore use this assumption in our main model in Sections 3-5. In Section 6 we consider concave and linear cases.

For technical convenience, we further assume that  $r(\theta)$  is strictly convex in  $\theta$ , and  $r'(\cdot), r''(\cdot)$  exist. These assumptions allow us to simplify the analysis. We now introduce a function:  $G(\theta) \equiv \theta r'(\theta) - r(\theta) - c$ , which we show later is the “virtual surplus” associated with type- $\theta$  customer when she is offered a specific version. Its structural properties are used in the subsequent analysis. Let  $\theta^* > 0$  denote the solution to  $G(\theta) = 0$ .

**Lemma 2.**  *$G(\theta)$  is strictly increasing for  $\theta > 0$ . Moreover,  $\forall c \geq 0$ ,  $\theta^*$  is unique.*

The above lemma implies that the potential revenue the seller can obtain from selling to a customer is increasing in the customer’s valuation. Moreover, given the variable cost  $c$ , a seller discards customers with valuation less than  $\theta^*$ .

---

<sup>3</sup>Note that we do not impose restrictions on the cost function  $s(q)$ , except that for all customers a maximizer exists (which makes  $r(\theta)$  well-defined). Note also that for any customer, the favorite option need not even be unique.

### 3 First-degree price discrimination

We first assume the seller can observe customers' types. This benchmark case is not only illustrative but facilitates why our model stands alone from all others in the existing literature. Following the technique of backward induction, we start with the production stage.

**Proposition 1.** *Let  $\bar{q}$  denote the quality level chosen in the development stage. Then if  $\bar{q} \leq r'(\theta^*)$ , then no customer is served. If  $\bar{q} > r'(\theta^*)$ , then for a given  $\bar{q}$ , there exists a unique pair  $(\underline{\theta}(\bar{q}), \bar{\theta}(\bar{q}))$  with  $\underline{\theta}(\bar{q}) < \theta^* < \bar{\theta}(\bar{q})$  such that the seller provides  $q^{FB}(\theta) = \bar{q}$  to customers with  $\theta \in [\underline{\theta}(\bar{q}), \bar{\theta}(\bar{q})]$  and no other customer purchases the product. Moreover,  $\forall \bar{q}_1, \bar{q}_2$  s.t.  $\bar{q}_1 \leq \bar{q}_2$ , we have  $\underline{\theta}(\bar{q}_1) \geq \underline{\theta}(\bar{q}_2)$  and  $\bar{\theta}(\bar{q}_1) \leq \bar{\theta}(\bar{q}_2)$ .*

Under first-degree price discrimination, every customer that is offered a version receives the same quality level, but is charged a different price. Whenever the transaction is efficient, trade always occurs, and the reservation utilities of those excluded customers are so high that the seller finds it unprofitable to even offer the highest possible quality. Moreover, as the quality limit  $\bar{q}$  increases, the set of customers served enlarges from both ends, and the low-end customers also benefit from the technology shift. In particular, when  $\bar{\theta}(\bar{q})$  hits the upper bound  $R$  of  $\theta$ 's support, the seller's incentive to increase the quality limit arises due to (1) the ability to charge a higher price for high-end customers; (2) the ability to include more low-end customers. Note also that when  $\bar{q} > r'(\theta^*)$ , the type- $\theta^*$  customer is always served under first-degree price discrimination.

The fact that every customer who is served receives the same quality level is in strict contrast with the majority of results in the nonlinear pricing literature. In that literature, it is common to assume the strict concavity of the social surplus  $s(q, \theta) \equiv u(q, \theta) - c(q)$ , see, e.g., Jullien [2000] and Sundararajan [2004]. With this assumption and the single-crossing condition ( $u_{q\theta}(q, \theta) > 0$ ,  $\forall q, \forall \theta$ ), we can show that the first-best quality level  $q^{FB}(\theta)$  is strictly increasing in  $\theta$ .<sup>4</sup> In our information good pricing framework, especially the software versioning scenario, this seems to be implausible, since it implies that some customers strictly prefer technologically inferior versions. If the price is not a concern, does a customer really prefer a student edition of Mathematica that cannot perform a huge number of functions/macros to the enterprise edition? Do people feel excited when they realize that

---

<sup>4</sup>The strict concavity implies that a unique solution  $q^{SB}(\theta)$  can be obtained from the first-order condition, i.e.,  $s_q(q^{FB}(\theta), \theta) = 0$ . Differentiating this equality by  $\theta$ , we have  $s_{qq}(q^{FB}(\theta), \theta) \frac{dq^{FB}(\theta)}{d\theta} + s_{q\theta}(q^{FB}(\theta), \theta) = 0$ . Note that  $s_{qq}(q^{FB}(\theta), \theta) < 0 < s_{q\theta}(q^{FB}(\theta), \theta)$ , we conclude that  $\frac{dq^{FB}(\theta)}{d\theta}$  must be strictly positive, and hence every served customer receives a version specific for her.

some functions of the software they just obtained are intentionally disabled? In this context, assuming every customer prefers the best quality makes better sense.

The next step is to consider the quality selection problem in the development stage.

**Theorem 1.** *Let  $\check{q}$  denote the unique solution to the equation  $\int_{\underline{\theta}(\check{q})}^{\bar{\theta}(\check{q})} \theta f(\theta) d\theta = E\theta$ , and  $\bar{q}^{FB}$  denote the optimal quality limit in the first-best scenario. Then  $\bar{q}^{FB}$  can be obtained by searching over points that satisfy  $\int_{\underline{\theta}(\bar{q})}^{\bar{\theta}(\bar{q})} \theta f(\theta) d\theta = C'(\bar{q})$ , provided that  $\bar{q} \geq r'(\theta^*)$ . If the above equation has no solution, then  $\bar{q}^{FB} = 0$ .*

*In particular, if  $C'(r'(\theta^*)) > E\theta$ , then  $\bar{q}^{FB} = 0$ ; if  $C'(\check{q}) > E\theta$ , then  $\bar{q}^{FB} < \check{q}$ . Moreover, in all cases  $\bar{q}^{FB} > r'(\theta^*)$ , and choosing any quality limit less than  $r'(\theta^*)$  is a strictly dominated strategy, independent of the structure of the development cost.*

This theorem characterizes the optimal level of quality limit in the first-best scenario, and has a clear economics intuition. Any choice below the critical level  $r'(\theta^*)$  is suboptimal since by offering it no transaction is efficient but the seller pays the development cost. If the development cost is fairly high (i.e., if  $C'(r'(\theta^*)) > E\theta$ ), then the seller finds it unprofitable to develop the information goods, and no transaction occurs due to the inefficiency. When the development cost is moderate, the optimal quality limit falls in the region  $[r'(\theta^*), \check{q}]$ .

## 4 Second-degree price discrimination

We now consider the optimal strategy to achieve second-degree price discrimination. We first take the quality limit  $\bar{q}$  as given, and derive the optimal quality-price schedule assuming that the seller offers versions to only an interval of customers. We next allow arbitrary exclusions of customers, and show that it is in the seller's best interest to serve only an interval of customers. Finally, we consider the optimal quality limit in the development problem.

We make the following assumption regarding the distribution of  $\theta$  in the sequel. Let  $F^c(\cdot) = 1 - F(\cdot)$  be the complementary cdf of  $\theta$ .

**Assumption 1.**  *$\theta F^c(\theta)$  is unimodal and has a unique maximum at  $k \in (0, R)$ .*

In particular, Assumption 1 implies that the function  $F^c(\theta) - \theta f(\theta)$  is initially positive and then becomes and stays negative.<sup>5</sup>

---

<sup>5</sup>This is more general than the regularity condition  $\frac{d}{d\theta} \frac{1-F(\theta)}{f(\theta)} < 0$  (the monotone hazard rate, or increasing failure rate (IFR) property). IFR is adopted in the screening literature to exclude the possibility of bunching. If a distribution is IFR, it also has the increasing generalized failure rate property (IGFR), namely,  $\theta f(\theta)/(1-F(\theta))$  is increasing in  $\theta$ . IGFR is a sufficient condition for Assumption 1, see Lariviere [2006].

If we interpret  $\theta$  as the price and the complementary cdf as the effective demand,  $\theta F^c(\theta)$  represents the

Before our analysis, let us first explain the intuition on how we obtain the optimal schedule. We first conjecture that at optimality the seller serves an interval of customers, and the seller should offer high-type customers higher quality levels to extract more revenue. The development cost of flagship product is a sunk cost, and hence it does not make sense not to sell the flagship product. Furthermore, the flagship product has to be offered to high-end customers among those customers that are served. This leads to a natural candidate of menu of versions. We will in the sequel formulate the optimization problem following this logic, and then verify its optimality.

#### 4.1 Optimal schedule when an interval of customers are served

We start with the case when the seller offers versions to an interval of customers. We will first take the interval as given and characterize the optimal quality-price schedule under such an assumption. We then allow the seller to choose one interval arbitrarily, and find the optimal boundary points that maximize the seller's profit.

We will assume that the seller offers a menu of versions to customers with  $\theta \in [\underline{\theta}, \tau)$ , customers with  $\theta \in [\tau, \bar{\theta}]$  accept the same version with quality limit  $\bar{q} \equiv q(\tau)$  and price  $p(\tau)$ , and customers in  $[0, \underline{\theta}) \cup (\bar{\theta}, R]$  are excluded, where  $0 \leq \underline{\theta} \leq \tau \leq \bar{\theta} \leq R$ . We further assume that by accepting  $(\bar{q}, p(\tau))$ , the type- $\tau$  customer receives her reservation utility, i.e.,  $p(\tau) = \tau\bar{q} - r(\tau)$ , independent of the quality-price schedule give for customers with  $\theta \in [\underline{\theta}, \tau)$ . We will verify later that this is a necessary condition for optimality.<sup>6</sup>

Suppose the customers with  $\theta \in [\underline{\theta}, \tau)$  are offered versions with  $(q(\theta), p(\theta))$  being the revenue as a function of price. If it is unimodal, there is an unambiguous revenue maximizing solution. Thus this provides a more natural interpretation than IFR.

<sup>6</sup>For a given quality limit  $\bar{q}$ , these thresholds  $\underline{\theta}, \tau, \bar{\theta}$  shall be functions of  $\bar{q}$ , but for notational ease we suppress this dependence in the analysis. Notice also that we do not exclude the possibilities of  $\underline{\theta} = 0$ ,  $\underline{\theta} = \tau$ ,  $\tau = \bar{\theta}$ , or  $\bar{\theta} = R$ , which represent respectively the cases when no low-end customer is excluded, no versioning occurs, only one customer receives the quality limit, and no high-end customer is excluded.

quality and price. The seller's problem is to find a quality-price schedule that solves :

$$\begin{aligned}
& \max_{q(\cdot), p(\cdot)} \left\{ (p(\tau) - c)(F(\bar{\theta}) - F(\tau)) + \int_{\underline{\theta}}^{\tau} (p(\theta) - c)f(\theta)d\theta \right\}, \\
& \text{s.t. (IC-1)} \quad \theta \in \operatorname{argmax}_{z \in [\underline{\theta}, \tau]} \theta q(z) - p(z), \quad \forall \theta \in [\underline{\theta}, \tau], \\
& \quad \text{(IC-2)} \quad \theta \bar{q} - p(\tau) \geq \max_{z \in [\underline{\theta}, \tau]} \theta q(z) - p(z), \quad \forall \theta \in [\tau, \bar{\theta}], \\
& \quad \text{(IC-3)} \quad r(\theta) \geq \max_{z \in [\underline{\theta}, \tau]} \theta q(z) - p(z), \quad \forall \theta \in [0, \underline{\theta}], \\
& \quad \text{(IC-4)} \quad r(\theta) \geq \max_{z \in [\underline{\theta}, \tau]} \theta q(z) - p(z), \quad \forall \theta \in [\bar{\theta}, R], \\
& \quad \text{(PC-1)} \quad \theta q(\theta) - p(\theta) - r(\theta) \geq 0, \quad \forall \theta \in [\underline{\theta}, \tau], \\
& \quad \text{(PC-2)} \quad \theta \bar{q} - p(\tau) \geq r(\theta), \quad \forall \theta \in [\tau, \bar{\theta}].
\end{aligned} \tag{1}$$

In Eq. (1), the first four inequalities are incentive compatibility (IC) conditions, where (IC-1) is for a customer that receives a version specific for herself, (IC-2) is for those customers that accept the same version with quality  $\bar{q}$ , and (IC-3) and (IC-4) are for respectively customers whose types are excluded from below and above. The last two inequalities in Eq. (1) represent participation constraints, i.e., each customer should get at least her reservation utility. Since customers with  $\theta \in [0, \underline{\theta}] \cup [\bar{\theta}, R]$  obtain their reservation utilities, their participation constraints are automatically satisfied. The optimal quality price schedule is summarized below, where  $\theta^*$  solves  $G(\theta^*) = 0$ .

**Theorem 2.** *Suppose that  $\bar{q}$  is given and the seller wishes to obtain second-degree price discrimination. Then customers with  $\theta \in [0, \theta^*)$  are not served, independent of  $\bar{q}$ . Transactions occur if and only if  $R > \theta^*$  and  $r'(\theta^*) < \bar{q}$ , in which case  $\exists \tau \in (\theta^*, k]$  and  $\bar{\theta}(\tau)$  such that*

- *Customers with  $\theta \in [0, \theta^*) \cup (\bar{\theta}(\tau), R]$  are not served.*
- *Each customer with  $\theta \in [\theta^*, \tau)$  receives a specific version with  $q(\theta) = r'(\theta)$ ,  $p(\theta) = \theta r'(\theta) - r(\theta)$ . No information rent is left for any customer in this region.*
- *Customers in  $[\tau, \bar{\theta}(\tau)]$  accept the same version with quality  $\bar{q}$  and price  $\tau \bar{q} - r(\tau)$ , and everybody in the interior of this region receives a nonzero surplus.*
- *The seller gets positive profit from every customer she serves.*
- *$\bar{\theta}(\tau) = R$  if  $r(R) \geq (R - \tau)\bar{q} + r(\tau)$ ; otherwise,  $r(\bar{\theta}(\tau)) = (\bar{\theta}(\tau) - \tau)\bar{q} + r(\tau)$ .*

- The value of  $\tau$  is then determined by the exhaustive search of local maxima on points in  $[0, k]$  that satisfy  $[\tau\bar{q} - r(\tau) - c]f(\bar{\theta}(\tau))\frac{d\bar{\theta}(\tau)}{d\tau} + (\bar{q} - r'(\tau))[F(\bar{\theta}(\tau)) - F(\tau) - \tau f(\tau)] = 0$ .<sup>7</sup>

When either  $\bar{q} \leq r'(\theta^*)$  or  $R \leq \theta^*$ , the seller is unable to make any profit by offering versions and maintaining customers' incentive compatibility, and therefore no transaction occurs. Transactions are efficient when  $\theta\bar{q} \geq c$ , but the information asymmetry drives out the possibility of transactions. If  $\bar{q} > r'(\theta^*)$  and  $R > \theta^*$ , the seller will offer different versions to customers. This is in contrast with the scenario where customers are endowed with a common reservation utility (Bhargava and Choudhary [2001]), where versioning is known to be suboptimal. Our result uncovers an incentive for the seller to provide different versions. As customers possess convex reservation utility, versioning helps the seller to extract more profits even if the customers possess constant marginal willingness to pay, and the products do not exhibit network effects. The inclination to provide versioning is fairly strong since the production cost does not change as a different quality level is provided.

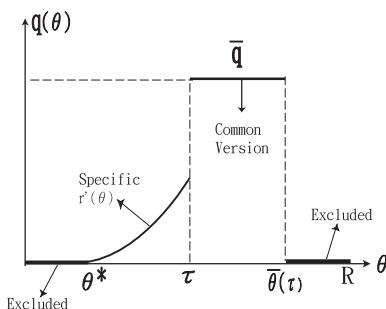


Figure 1: An example of the optimal quality schedule under the second-degree price discrimination.

Furthermore, Theorem 2 characterizes the optimal quality-price schedule, whereas a generic shape of the quality levels offered to customers is presented in Fig. 1. At optimality, the seller discards both the low-end and high-end customers. For those served, the seller extracts full information rent for customers with relatively low willingness to pay, and offers a common version to the rest.<sup>8</sup>

<sup>7</sup>If  $f(\cdot)$  is widely spread-out, i.e.,  $f(\bar{\theta}(\tau))$  is relatively small, the second term  $(\bar{q} - r'(\tau))[F(\bar{\theta}(\tau)) - F(\tau) - \tau f(\tau)]$  dominates. Since  $\bar{q} - r'(\tau) > 0$ , the sign of the derivative depends only on  $F(\bar{\theta}(\tau)) - F(\tau) - \tau f(\tau)$ , which approximately is the derivative of  $\tau(1 - F(\tau))$  as  $\bar{\theta} \rightarrow R$ , and hence it turns negative right at  $\tau = k$ . This ties in with the unimodality assumption (Assumption 1).

<sup>8</sup>The rationale to exclude low-end customers is clear in the standard nonlinear pricing literature: the seller is unable to extract positive rent from a low-end customer, and therefore the seller should not serve her. The cutoff point is one at which the virtual surplus turns positive.

In the region  $[\theta^*, \tau)$ , each customer is offered a version specific for her. Nevertheless, by accepting it the customer receives exactly her reservation utility. The seller is able to fully extract the information rent from customers in this region, but has to distort the quality levels away from the first-best levels to maintain incentive compatibility. Inefficiency occurs due to this, because the seller cannot observe customers' types. Moreover, the offered quality level is strictly increasing in the type, i.e., customers with higher marginal willingness to pay receive products of better quality.

The customers in  $[\tau, \bar{\theta}(\tau)]$  are offered a common version that makes the type- $\tau$  customer receives her reservation utility. The seller earns no information rent from these customers. The upper bound of this region is determined by the critical customer who is indifferent to accepting this version and staying with her outside opportunity if such a critical customer exists; otherwise, the upper bound is  $R$ , i.e., no high-end customer is excluded. Moreover, at the critical point  $\tau$  we see a clear discontinuity in the version specification: both the quality and the price have jumps at  $\theta = \tau$ . This implies that the profit the seller collects is also discontinuous at  $\tau$  since the marginal cost is independent of the quality.

Finally, the exclusion of high-end customers is due to the pre-determined quality limit in the development stage, and hence with the linear utility format and strictly convex reservation utility, the seller must give up those high-end customers because their outside opportunities are too high. This was observed in the first-best scenario, even though the determination of the cutoff type is based on a different criterion.

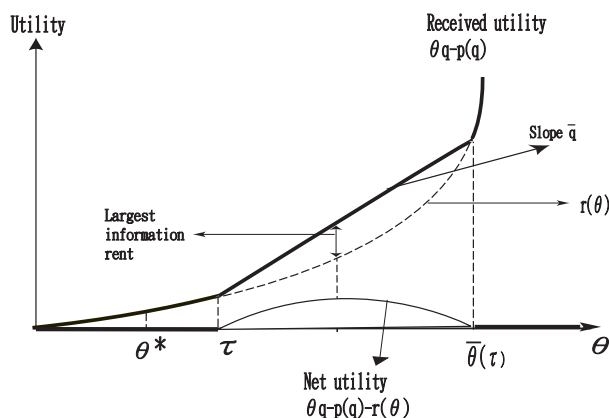


Figure 2: An example to show the net utility under the second-degree price discrimination.

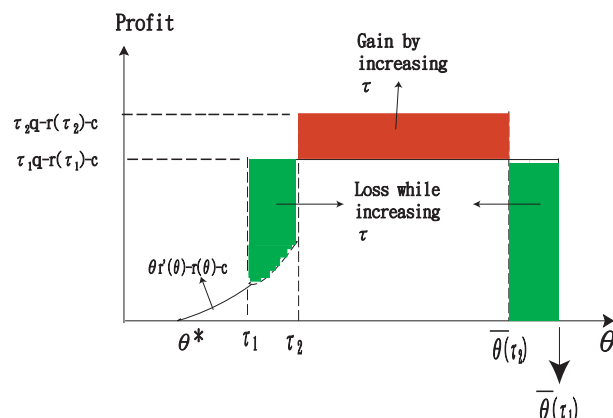


Figure 3: The influence of profits while changing the value of  $\tau$ .

The shape of net utilities is also worth noting. Assuming that some high-end customers



are excluded by the optimal quality-price schedule, we draw in Fig. 2 the received utilities of customers and the “net” utilities (the received utility net the reservation utility). Except  $\theta \in (\tau, \bar{\theta}(\tau))$ , customers receive their reservation utilities in the end, regardless of whether she accepts a version or stays unserved. Inside the region  $(\tau, \bar{\theta}(\tau))$ , the net utility is unimodal, and the customer that receives the maximal rent is located in the interior.

The optimal schedule can be interpreted as follows. Sundararajan [2004] reports that the optimal schedule can be decomposed into two parts: one is driven by the outside opportunity and the other is determined by the heterogeneity of utility. In our model, the first one corresponds to the case  $\theta \in [\theta^*, \tau)$ , where the optimal quality schedule is determined by customers’ outside opportunities. Nevertheless, the seller obtains positive profits from these customers, in contrast to Jullien [2000]. The second part corresponds to the region  $\theta \in [\tau, \bar{\theta}]$ . If  $r(\cdot)$  were constant, a common version with  $\bar{q}$  would be offered to all customers but at a lower price. Thus, we see the price shift alluded to above (c.f. Sundararajan [2004]).

Fig. 3 demonstrates the trade-off the seller faces while choosing the value of  $\tau$ . As a seller increases  $\tau$  from  $\tau_1$  to  $\tau_2$ , i.e., she increases the starting point of offering a common version, the price for that common version increases even though the quality remains  $\bar{q}$ . This change influences the profit in three ways: First, the seller loses some profits on customers with  $\theta$  between  $\tau_1$  and  $\tau_2$ , since the quality levels offered to them are  $\{r'(\theta)\}$ ’s rather than  $\bar{q}$ . Second, the shift of  $\tau$  increases the profit gained from those who accept the common version; the seller gets  $(\tau_2 - \tau_1)\bar{q} - (r(\tau_2) - r(\tau_1))$  more in the region  $[\tau_2, \bar{\theta}(\tau_2)]$ . Third, since for a given quality limit  $\bar{q}$  this version is priced higher (from  $\tau_1\bar{q} - r(\tau_1)$  to  $\tau_2\bar{q} - r(\tau_2)$ ), fewer customers are willing to purchase, and hence this shift excludes more high-end customers. The optimal value of  $\tau$  balances the gains and losses.

## 4.2 Optimal schedule and target quality with arbitrary exclusion

Theorem 2 provides the optimal quality-price schedule when an interval of customers is served. An immediate question is whether this schedule remains optimal if the seller can exclude customers arbitrarily, e.g., she excludes customers with  $\theta \in [0.23, 0.31] \cup [0.45, 0.79]$ . The set of excluded customers can be even more sophisticated, namely any measurable set with respect to the probability space  $([0, R], \mathcal{B}, F(\cdot))$ , where  $\mathcal{B}$  is the collection of Borel measurable sets over  $[0, R]$ .<sup>9</sup> Nevertheless, our proposed quality schedule is indeed optimal.

---

<sup>9</sup>The seller might want to create two products that are far apart so that the price effect does not lead to cannibalization. Separation of quality levels could also help the seller to manage them easily and change them independently. Thus, we have to examine this possibility before arriving at a conclusion regarding the optimality of the quality schedule.

**Theorem 3.** *Given the quality limit  $\bar{q}$ , the quality-price schedule proposed in Theorem 2 is optimal even if arbitrary exclusion is allowed.*

The proof follows the approach of Jullien [2000]. We show that no intermediate exclusion is profitable, and hence at optimality the seller must offer versions to an interval of customers. Since the schedule proposed in Theorem 2 is optimal if customers in an interval are served, its optimality continues to hold in this broader class of schedules. This implies that the optimal versioning has a specific structure, namely that, the seller offers a pool of customers a common version and excludes some customers. This is labelled as “bunching with exclusion” in the literature, and it appears due to the unique cost structure of information goods.

Our characterization and verification of optimal quality-price schedule is now complete. The seller’s problem in the development stage is as follows. Let  $V(\bar{q})$  be the optimal value of Eq. (1) when quality-price schedule is optimally chosen. The optimal quality limit can be found through exhaustive search of the local maxima:  $\bar{q}^{SB} = \operatorname{argmax}_{\bar{q}}\{V(\bar{q}) - C(\bar{q})\}$ .

## 5 Comparative statics

In this section we discuss the comparative statics of our model. This includes (1) given a fixed quality limit, how does second-degree price discrimination differ from the first-degree price discrimination? (2) How does the profit change as the quality limit varies? We first compare these two informational scenarios while assuming a fixed quality limit  $\bar{q}$ .

**Theorem 4.** *Suppose that the quality limit  $\bar{q}$  is given. Then*

- *If  $\bar{q} > r'(\theta^*)$ , let  $\underline{\theta}^{FB}(\bar{q}), \underline{\theta}^{SB}(\bar{q}), \bar{\theta}^{FB}(\bar{q}), \bar{\theta}^{SB}(\bar{q})$  denote respectively the lowest and highest type of customers that are offered a version under the two price discrimination. Then for all  $\bar{q}$ ,  $\underline{\theta}^{FB}(\bar{q}) < \underline{\theta}^{SB}(\bar{q}) < \bar{\theta}^{SB}(\bar{q}) < \bar{\theta}^{FB}(\bar{q})$ . In particular, customers with  $\theta \leq \theta^*$  are never served under second-degree price discrimination, whereas an interval around the type- $\theta^*$  customer is included in first-degree price discrimination.*
- *Under first-degree price discrimination, each customer either is not served or receives  $\bar{q}$ . However, under the second-degree price discrimination, a continuum of versions may be offered. If at optimality the seller chooses  $\tau = (r')^{-1}(\bar{q})$ , then only the customer with  $\theta = (r')^{-1}(\bar{q})$  receives the efficient quality level  $\bar{q}$ .*
- *When  $R \leq \theta^*$ , where  $\theta^*$  is the critical customer whose virtual surplus just turns positive, under the second-degree price discrimination the seller will not develop the information*

*goods, regardless of the cost structure  $C(\cdot)$ . Nevertheless, there exist situations where the seller does use first-degree discrimination over customers.*

The first comparison shows that the set of customers offered under the second-degree price discrimination is a proper subset of that under the first-degree price discrimination. The information asymmetry does prevent the seller from serving some customers although the transactions are efficient. The second comparison demonstrates the inefficiency on the quality levels offered under the second-degree price discrimination. Except possibly a subset of customers, a continuum of customers receive versions that have inferior quality levels. By intentionally shading the quality levels, the seller gains against the information asymmetry.

Finally, to induce development of information goods, the second-best scenario requires a larger maximal marginal willingness to pay of customers. If we interpret value of  $R$  as a measure of customers' heterogeneity, a higher degree of heterogeneity among customers is needed to overcome the information asymmetry faced by the seller.

Now we discuss the impact of different quality limits on the quality-price schedule under both first- and second-degree price discrimination.

**Theorem 5.** *Suppose the two quality limits  $\bar{q}_1, \bar{q}_2$  are pre-determined, and  $\bar{q}_2 > \bar{q}_1$ . Then in the production stage, the seller obtains strictly greater profits with  $\bar{q}_2$  compared to the case with  $\bar{q}_1$  under both first- and second-degree price discrimination.*

In the first-best scenario, when a higher quality limit is set in the development stage, the prices are higher and the set of served customers is larger. Under second-degree price discrimination, as a higher quality limit is chosen, the seller can always choose the same starting point of offering a common version. By doing so she gains in two ways: (1) the price of this common version is strictly higher; (2) more high-end customers are willing to purchase this version compared to the case with  $\bar{q}_1$ . This simple rule is particularly useful if a sudden increase of quality limit takes place, since the seller can increase her profit by introducing a new flagship product without changing the extant versions.

## 6 Discussions and extensions

In this section we first extend our model to incorporate more general variable costs, customers' utilities, and reservation values. We then discuss managerial implications.

## 6.1 Extensions

We first let the reservation utility be either linear or concave rather than convex, and allow the variable cost  $c(q)$  to be linear or concave.<sup>10</sup> Observe that when the cost is convex in quality, even if customers do not have heterogeneous outside opportunities, it may be optimal for the seller to offer multiple versions. The case with convex cost function can be analyzed following the standard approach (e.g., Jullien [2000] and Mussa and Rosen [1978]), and therefore the detailed derivations are omitted.

We show that when  $r(\theta)$  is either linear or concave and  $c(q)$  is linear or concave, offering a single version is optimal.

**Theorem 6.** *Suppose that  $r(\theta)$  is concave or linear. Offering a single version is optimal if: (1) the distribution cost is independent of quality. (2)  $c(q)$  is linear or concave in  $q$  and the usual regularity condition  $\frac{d}{d\theta} \frac{1-F(\theta)}{f(\theta)} \leq 0 \leq \frac{d}{d\theta} \frac{F(\theta)}{f(\theta)}$  holds.*

The intuition for Theorem 6 is as follows. When the variable cost is constant or weakly concave, the only reason for the seller to offer multiple versions is to match the reservation utilities of a continuum of customers. However, when the reservation utility is linear or concave, the corresponding quality schedule that matches their outside options  $q(\theta) = r'(\theta)$  is decreasing in type. This would inevitably violate incentive compatibility. Linear or concave reservation utilities could arise if switching can be done relatively easily by high-end users, such as in the case of word processing or browsing the web.

Thus, we have established that when customers have constant marginal willingness to pay and the variable cost is constant or concave, the convexity of reservation utility is necessary to induce versioning. The next question we address is how changing the utility function and variable cost but keeping the reservation utility convex affect the structure of the optimal quality schedule. We examine this numerically. Specifically, we solve a mathematical program in which we incorporate nonlinear utility functions, and concave variable costs, but the reservation utilities remain convex. We focus on how these factors determine the number of versions offered and the exclusion of customers.

We first set up the optimization problem below. Let the utility function  $u(q, \theta)$  be  $\theta q - \rho \theta^2$ , where  $\rho \geq 0$ . Note that if  $\rho = 0$ , this degenerates to the linear utility case. We discretize the customers' types into  $n$  classes with corresponding weights  $\{w_i, \forall i = 1, \dots, n\}$ .

---

<sup>10</sup>As argued by Bhargava and Choudhary [2001], “[t]he costs of duplicating a higher-quality information good are not likely to be much higher than that for a low-quality product.”

The optimization problem is as follows.

$$\begin{aligned}
& \max_{\{\pi_i, y_i, q_i, p_i\}} \sum_{i=1}^n \pi_i w_i \\
\theta_i q_i - \rho \theta_i^2 - p_i & \geq \theta_i q_j - \rho \theta_i^2 - (1 - y_i)M, \forall j \neq i, \forall i, j = 1, \dots, n, \\
\theta_i q_i - \rho \theta_i^2 - p_i & \geq r(\theta_i) - (1 - y_i)M, \forall i = 1, \dots, n, \\
q_i & \leq \bar{q}, \forall i = 1, \dots, n, \\
\pi_i & \leq \min \{y_i M, w_i(p_i - c(q_i))\}, \forall i = 1, \dots, n,
\end{aligned}$$

where  $y_i \in \{0, 1\}, \forall i = 1, \dots, n$ , is the indicator of whether type- $i$  customer is included ( $y_i = 1$  if included), and  $M$  is a sufficiently large number. The first two constraints are (IC) and (PC) (where the extra  $(1 - y_i)M$  allows us to verify these two conditions for only those customers that are served). The third constraint depicts that the quality should not exceed the limit. The last one ensures that the profit earned from an excluded customer is zero and that from a served customer coincides with the revenue minus the variable cost at optimality.

Having formulated the optimization problem, we can now investigate a variety of scenarios. We consider situations when products are more profitable and those in which they are less profitable, and when the quality limit is high or moderate. We also examine how the degree of heterogeneity affects the optimal schedule. For this purpose, we assume  $\theta$  follows a uniform distribution over a bounded support, i.e.,  $w_i = \frac{1}{n}, i = 1, \dots, n$ . We represent the reservation utility by a simple power function  $r(\theta) = \xi \theta^a$ , where  $a > 1$  captures the curvature of customers' reservation values. The variable cost is chosen as  $c(q) = Kq^b$ , where  $K \geq 0$  represents how costly it is to offer the products (or how profitable the product is), and  $b \in [0, 1]$  measures the curvature of variable cost. The impact of the nonlinearity of utility on the optimal quality schedule is measured by  $\rho$ . The benchmark case has the following parameters:  $\rho = 0, \bar{q} = 3, a = 2, \xi = 0.4, K = 0.1$ , and  $b = 0.8$ . In the experiments we vary one parameter at a time and keep all other parameters the same to examine how sensitive the optimal quality schedule is with regard to each factor. We use LINGO (Lindo Systems Inc.) to solve the optimization problem.

In Fig. 4, we vary the value of  $K$ , and observe that reducing profitability may reduce versioning. In order to compensate the costly production/distribution of information goods, the seller should provide only the flagship product to extract more revenue. We also observe that when the product is less profitable, more low-end customers get excluded due to their low willingness to pay. In Fig. 5, we change the quality limit. Fig. 5 shows that increasing the quality limit might induce versioning, since the seller can choose quality over a wider

range. When the quality limit is high, the seller tends to offer multiple versions and cover all high-end customers. When the quality limit becomes moderate, the seller offers only a single version and gives up high-end customers due to their high reservation utilities. The exclusion of low-end customers is insensitive to the quality limit.

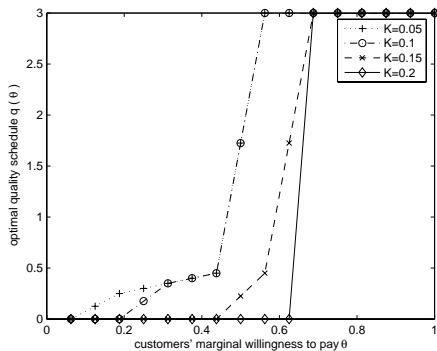


Figure 4: The quality schedule under different profitability.

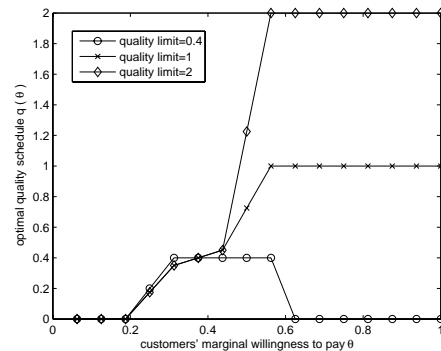


Figure 5: The quality schedule under different quality limits.

Figs. 6 and 7 demonstrate the effect of heterogeneity on versioning. In Fig. 6, we vary the exponent  $a$  of the reservation utility. We observe that when the reservation utility is highly heterogeneous (i.e.,  $a$  is large), the seller should offer more versions to carefully match outside options, and offer the flagship product to fewer customers. When  $a$  is very small ( $a = 1.1$ ), only one version is offered. In Fig. 7 we change the exponent  $b$  of the cost, and we find that only two kinds of quality schedules are possible. Note that there are only two regions: a full rent extraction region and the region where the quality limit is offered. Quality schedules in both regions are independent of the variable cost. Hence, when we change  $b$ , only the threshold between these two regions change. Moreover, when the variable cost is more concave, the seller excludes more low-end customers and offers fewer versions.

Finally, we introduce nonlinearity of utility by using positive  $\rho$ 's in Fig. 8 with low quality limit ( $\bar{q} = 1$ ) and Fig. 9 with high quality limit ( $\bar{q} = 3$ ). We observe that the outcome depends on the quality limit. The more nonlinear the utility, the more likely the seller is to exclude high-end customers. Nevertheless, whether nonlinearity induces versioning is ambiguous. Nonlinearity leads to fewer versions under low quality limit (as in Fig. 8) but more versions under high quality limit (Fig. 9). Whether the nonlinearity of utility affects the exclusion of low-end customers is also ambiguous. This might be because when we increase  $\rho$  to add more nonlinearity, the utility also becomes lower. The exclusion of high-end customers in

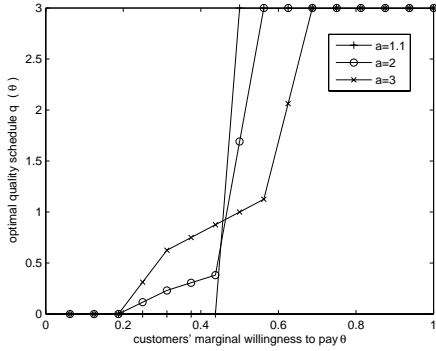


Figure 6: The quality schedule under different reservation utilities.

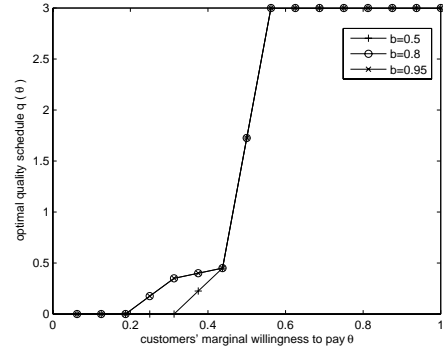


Figure 7: The quality schedule under different variable costs.

Fig. 8 also results from the reduction in utility. When  $\rho$  is high, the utility becomes lower, but the reservation utility is still the same.

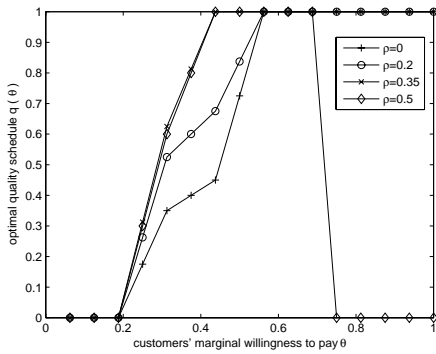


Figure 8: The quality schedule under different customers' utilities with  $\bar{q} = 1$ .

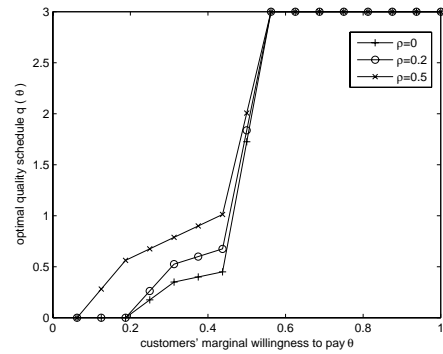


Figure 9: The quality schedule under different customers' utilities with  $\bar{q} = 3$ .

Table 2 summarizes our findings via these simulations.<sup>11</sup> In this table, we show the effects of increasing various factors on the number of versions, and exclusion of low-end and high-end customers. The variation in the optimal quality schedule due to the change in these factors is not significantly different from the benchmark case, except due to the change in quality limit. Therefore, we conclude that solving the two-stage problem is important in order to determine the quality limit.<sup>12</sup>

<sup>11</sup>We have tried many combinations of parameters and found that these insensitive results are robust against choices of parameters.

<sup>12</sup>Originally, we believed that exclusion of intermediate customers might be optimal for concave variable

Table 2: Summary of impacts of relevant factors on versioning

| Factor                   | No. of versions | exclusion of low-end | exclusion of high-end |
|--------------------------|-----------------|----------------------|-----------------------|
| profitability            | increases       | more likely          | insensitive           |
| quality limit            | increases       | insensitive          | less likely           |
| convexity of $r(\theta)$ | increases       | less likely          | insensitive           |
| concavity of $c(q)$      | decreases       | more likely          | insensitive           |
| nonlinearity of utility  | ambiguous       | ambiguous            | ambiguous             |

## 6.2 Managerial implications

Our analysis identifies a previously ignored driving force that induces versioning. If there are multiple effective outside opportunities that are accessible to target customers, then the seller should adopt versioning to extract more revenue; otherwise, she should offer a single version (illustrated in Fig. 6). To this end, careful investigation of available outside options facing target customers is critical to crafting a versioning strategy. The inherent difference of customers' outside opportunities could potentially explain why amongst software producers, some sell a single quality to all customers but others offer multiple versions, even though they have similar cost structures for product development and they all exhibit network effects. This could be due to lack of multiplicity of outside options.

Our results also suggest that when customers have access to various outside options, it may be suboptimal for the seller of information goods to always serve all high-end customers. Even though high-end customers have higher willingness to pay, they are also endowed with higher outside opportunities. For example, Decisioneering, Inc. offers multiple versions of Crystal Ball and serves only the mid-range customers (high-end customers are captured by Computer Aided Design of Microelectronic Packages or ASTi FAA Commercial Simulation Package, and low-end customers use Excel). Another example is Microsoft Encarta. Encarta currently has Academic, Premium, and Standard versions.<sup>13</sup> However, customers that need to look up specific information regularly may prefer Britannica, and low-end customers can use free online encyclopedia such as Wikipedia and would not purchase Encarta. To extract more revenue, the seller has to fine-tune the quality levels of versions (below the high-end) so that these offers match customers' outside options. Sufficient knowledge of customers'

---

costs and nonlinear utility functions. However, we are unable to uncover this phenomenon in our numerical experiments. We conjecture that the structure of optimal quality schedule (Fig. 1) might occur in most scenarios that are of practical interest.

<sup>13</sup><http://www.microsoft.com/products/encarta/> (last cited on January 14, 2007).



outside options is essential for doing this fine-tuning (Figs. 4, 6, and 7). This might be an issue for Bloomberg and SAP.

In order to implement the optimal schedule summarized in Theorem 2, two thresholds need to be determined, namely, the exclusion point of low-end customers and the starting point at which to offer the flagship product. The exclusion point of low-end customers is completely determined by the reservation utility and distribution cost, independent of the quality limit and the distribution of customers' types. In contrast, the starting point of offering the flagship product is jointly determined by all the above inputs. This demonstrates the different degrees of sophistication needed in order to locate these two key thresholds. For software, the basic editions are usually easy to design. In contrast, there are far more issues to consider when positioning professional editions, such as functions of user interface, image resolution, speed of operation, and compatibility (e.g., Maple, Microsoft Money, TurboTax). For information services, designing appropriate high-end products to avoid cannibalization is also a difficult task, compared to the basic or free-sponsored versions (Classmates.com, CNN, NY Times, Yahoo!). The sellers have to carefully investigate customers' time values, technical background, heterogeneity of willingness to pay, and the largest benefit their current technology can provide. This is more critical when the quality limit is low (as seen in Fig. 5).

Sometimes marketing departments manage to acquire more information from target customers so that certain degree of personalization/customization is possible, but the flagship product cannot be changed immediately (it cannot be achieved without going back to the development phase). We suggest that in this case, they could enlarge the base of served customers from both ends. Note that under first-degree discrimination only the flagship product is offered. Thus, when managers have more detailed information regarding customer preferences, they should reduce the span of versioning (e.g., reduce the number of versions, narrow down the quality differences), and focus on better horizontal differentiation to achieve customization. This is the case for Virtual Vineyards: they regularly make special offers to customers based on their clickstreams. Another example of personalized pricing is the market for computer servers. Hewlett Packard, IBM, and Sun Microsystems conduct an ROI (Return on Investment) analysis of their customer accounts. They then offer various personalized discounts over identical products to customers based on their ROIs. LexisNexis charges libraries different prices even though they offer them identical information service.

Finally, if a sudden increase of quality limit takes place, but redesigning the versioning strategy either is time-consuming or could cause dissatisfaction amongst existing customers, using the proof of Theorem 5 we can construct a simple rule to increase the revenue in-

Table 3: Comparison of our paper with previous literature on versioning information goods

| Paper                         | Utility   | Variable cost | Reservation utility | Versions | Exclusion |
|-------------------------------|-----------|---------------|---------------------|----------|-----------|
| Bhargava and Choudhary [2001] | linear    | concave       | homogeneous         | one      | low end   |
| Bhargava and Choudhary [2004] | linear    | 0             | homogeneous         | two      | low end   |
| Ghose and Sundararajan [2005] | quadratic | 0             | homogeneous         | multiple | low end   |
| Jing [2006]                   | linear    | 0             | homogeneous         | two      | low end   |
| Jones and Mendelson [1998]    | linear    | 0             | homogeneous         | one      | low end   |
| Raghunathan [2000]            | quadratic | 0             | homogeneous         | multiple | N/A       |
| This paper                    | both      | concave       | convex              | multiple | both ends |
|                               | both      | concave       | linear/concave      | single   | low end   |
|                               | both      | convex        | convex              | multiple | both ends |
|                               | both      | convex        | linear/concave      | multiple | low end   |

stantaneously. The seller can introduce a new flagship product, keep extant versions, and modify the pricing schedule to avoid cannibalization. An unexpected increase of quality limit may be due to a significant technological advance made by internal R&D groups, the acquisition of more advanced modules that improve the quality, release of features that were kept secret for strategic concerns, etc. This issue is particularly important for the software industry. Due to its rapid and evolving technological improvement, many software packages have successive generations and multiple versions (e.g., Acrobat, Encyclopaedia Britannica, Illustrator, and McAfee). This rule may be useful to determine the pricing of these software generations and versions.

## 7 Conclusion

In this paper we consider a two-stage problem for information goods production. We include all ingredients of information goods and provide novel guidelines to sellers that wish to adopt a versioning strategy. The deviations of our model setting and results are listed in Table 3. We show that versioning is profitable when customers possess convex reservation utilities, and characterize the optimal quality-price schedule. In the optimal strategy, the seller discards both the low-end and high-end customers. For those served, the seller extracts full information rent from customers with relatively low willingness to pay, but offers a common version to the rest. We also provide a simple rule for selecting the optimal quality limit in both cases, and perform comparative statics under both types of discrimination. We then extend our model to incorporate concave variable costs and concave reservation utilities. We find that versioning is suboptimal if the reservation utility is linear or concave, but it might occur at optimality with convex reservation utilities. Thus, the profitability of

versioning depends significantly on the structure of customers' reservation utilities.

Several extensions arise naturally, a particular one being the incorporation of network externalities. When network effects are measured by the total usage of the product across different versions, the seller always has an incentive to include more customers. However, if products of different quality levels share only part of the benefit, the seller faces an intriguing trade-off: should she offer user-specific versions to fully extract low-end customers' rent or should she offer only a limited number of versions to induce higher network effects?

Also, most software products have different maintenance programs. As pointed out in Shapiro and Varian [1998], the maintenance of software products is very costly, and it critically depends on the quality the company offers to customers. This trade-off is worth investigating. Another direction that seems worth while to pursue is to extend the analysis to a dynamic setting in which generations of customers purchase the product and the distribution of reservation utilities varies over time. Since the choice of quality limit is irreversible, the seller faces a constrained optimization problem in the development stage with respect to the current quality limit. Because developing a new (and higher) quality limit is costly, the dynamic setting may allow us to predict the optimal timing for investing in new product development in such an industry.

Introducing competition amongst sellers is a potential avenue for research. As sellers choose quality limits upfront, they may distinguish themselves by selecting different levels, and therefore adopt different quality-price schedules given the quality limits. This differentiation bypasses the head-to-head price competition that could potentially drive away all the profit. Extending our research to incorporate quantity differentiation is another fruitful direction to pursue. The seller's problem can then be regarded as one in which customers obtain more utility from possessing or consuming more units of the information good. This introduces the scope for enhancing profitability through quantity differentiation, a strategy that is commonly adopted by mainframe application providers (e.g., IBM) and manufacturers of enterprise database systems such as Oracle.

## Appendix. Proofs

### Proof of Lemma 2

Differentiating  $G(\theta)$  by  $\theta$ , we have  $G'(\theta) = \theta r''(\theta) \geq 0$ , and hence  $G(\theta)$  is increasing and  $G(\theta) = G(0) + \int_0^\theta x r''(x) dx$ . By strict convexity,  $r'(\theta) \rightarrow \infty$  as  $\theta \rightarrow \infty$ . Therefore,  $\lim_{b \rightarrow \infty} \int_0^b r''(\theta) d\theta = \infty$ . Let  $M$  be an arbitrary large number. Since  $\lim_{b \rightarrow \infty} \int_0^b r''(\theta) d\theta = \infty$ , given any constant  $C_1$ ,  $\forall M_1 \equiv C_1 \times M$ , there exists another constant  $C_2$  such that

$\int_{C_1}^{C_2} r''(\theta)d\theta > M$ . Multiplying the integrand by  $\theta$ , we obtain that  $\int_{C_1}^{C_2} \theta r''(\theta)d\theta > M$ , and therefore  $\lim_{\theta \rightarrow \infty} G(\theta) = \lim_{b \rightarrow \infty} \int_0^b \theta r''(\theta)d\theta = \infty$ .

Since  $G(0) = -r(0) - c \leq 0$  and  $\lim_{\theta \rightarrow \infty} G(\theta) = \infty$ , a solution exists for  $\theta r'(\theta) - r(\theta) - c = 0$  from the intermediate value theorem, and it is unique by strict monotonicity of  $G(\theta)$ .  $\square$

### Proof of Proposition 1

If a customer is served, the seller's best response is to offer her  $\bar{q}$  due to the common marginal cost. Thus the maximum rent that the seller can extract from a type- $\theta$  customer is  $\theta\bar{q} - r(\theta)$ . Thus, the type- $\theta$  customer is served if and only if  $\theta\bar{q} - r(\theta) - c \geq 0$ .

Consider the case  $\bar{q} = r'(\theta^*)$ . In this case,  $\theta^*$  is a solution to  $\theta\bar{q} - r(\theta) - c = 0$ . Define  $H(\theta, \bar{q}) = \theta\bar{q} - r(\theta) - c$ . Differentiating  $H(\theta, \bar{q})$  by  $\theta$  while  $\bar{q} = r'(\theta^*)$ , we have  $\partial H(\theta, r'(\theta^*)) / \partial \theta = r'(\theta^*) - r'(\theta)$ , which is negative when  $\theta < \theta^*$  and positive when  $\theta > \theta^*$ . Thus  $H(\theta, r'(\theta^*))$  attains its maximum uniquely at  $\theta = \theta^*$ . This also implies that  $\theta\bar{q} - r(\theta) - c$  is negative for all  $\theta$ , and no customer shall be served. When  $\bar{q} < r'(\theta^*)$ , we have  $\theta\bar{q} - r(\theta) - c < \theta r'(\theta^*) - r(\theta) - c \leq 0, \forall \theta \geq 0$ . Thus, the seller serves no customer.

When  $\bar{q} > r'(\theta^*)$ ,  $\theta^*\bar{q} - r(\theta^*) - c > \theta^*r'(\theta^*) - r(\theta^*) - c = 0$ .  $H(0, \bar{q}) = -r(0) - c < 0, \forall \bar{q}$ , and thus by intermediate value theorem, there exists  $\underline{\theta}(\bar{q}) \in [0, r'(\theta^*))$  such that  $H(\underline{\theta}(\bar{q}), \bar{q}) = 0$ . From  $\partial H(\theta, \bar{q}) / \partial \theta = \bar{q} - r'(\theta)$ , eventually  $H(\theta, \bar{q})$  will become negative when  $\theta$  is sufficiently large. Therefore, there exists a constant  $\tau(\bar{q}) > \theta^*$  such that  $H(\tau(\bar{q}), \bar{q}) = 0$ . To see that  $H(\theta, \bar{q}) \geq 0$  if and only if  $\theta \in [\underline{\theta}(\bar{q}), \tau(\bar{q})]$ , since  $r'(\theta)$  is strictly increasing in  $\theta$  and  $H(\underline{\theta}(\bar{q}), \bar{q}) = H(\tau(\bar{q}), \bar{q}) = 0, r'(\theta) < \bar{q}$  if and only if  $\theta < \tau(\bar{q})$ . Thus if  $\theta > \tau(\bar{q})$ ,  $H(\theta, \bar{q}) = \theta\bar{q} - r(\theta) - c < \tau(\bar{q})\bar{q} - r(\tau(\bar{q})) - c - (\theta - \tau(\bar{q}))(\bar{q} - r'(\tau(\bar{q}))) = H(\tau(\bar{q}), \bar{q}) = 0$ , where the strict inequality follows from the strict concavity of  $H(\theta, \bar{q})$  with respect to  $\theta$ . The strict concavity of  $H(\theta, \bar{q})$  also implies that  $\forall \theta < \underline{\theta}(\bar{q})$ ,  $H(\theta, \bar{q}) = \theta\bar{q} - r(\theta) - c > \underline{\theta}(\bar{q})\bar{q} - r(\underline{\theta}(\bar{q})) - c - (\theta - \underline{\theta}(\bar{q}))(\bar{q} - r'(\underline{\theta}(\bar{q}))) > H(\tau(\bar{q}), \bar{q}) = 0$ , where we have used  $\bar{q} > r'(\underline{\theta}(\bar{q}))$  in the second inequality. A similar argument shows that  $H(\theta, \bar{q}) \geq 0, \forall \theta \in [\underline{\theta}(\bar{q}), \tau(\bar{q})]$ . Hence only customers with  $\theta \in [\underline{\theta}(\bar{q}), \tau(\bar{q})]$  are served.  $\square$

### Proof of Theorem 1

The optimal quality limit  $\bar{q}^{FB}$  solves  $\max_{\bar{q}} \{ \int_{\underline{\theta}(\bar{q})}^{\bar{\theta}(\bar{q})} [\theta\bar{q} - r(\theta) - c] f(\theta) d\theta - C(\bar{q}) : \bar{q} \geq r'(\theta^*) \}$ , where  $\underline{\theta}(\bar{q})$  and  $\bar{\theta}(\bar{q})$  are the two roots of  $\theta\bar{q} - r(\theta) - c = 0$ . Note that we have ignored the trivial case where  $\bar{q} \in (0, r'(\theta^*)]$ , since by doing so the seller gets a strictly negative payoff. Let  $\Pi^{FB}(\bar{q}) \equiv \int_{\underline{\theta}(\bar{q})}^{\bar{\theta}(\bar{q})} [\theta\bar{q} - r(\theta) - c] f(\theta) d\theta - C(\bar{q})$  denote the expected payoff when  $\bar{q}$  is chosen. Differentiating  $\Pi^{FB}(\bar{q})$ , we can express  $\frac{d\Pi^{FB}(\bar{q})}{d\bar{q}}$  as

$$[\bar{\theta}(\bar{q})\bar{q} - r(\bar{\theta}(\bar{q})) - c]f(\bar{\theta}(\bar{q}))\bar{\theta}'(\bar{q}) - [\underline{\theta}(\bar{q})\bar{q} - r(\underline{\theta}(\bar{q})) - c]f(\underline{\theta}(\bar{q}))\underline{\theta}'(\bar{q}) + \int_{\underline{\theta}(\bar{q})}^{\bar{\theta}(\bar{q})} \theta f(\theta)d\theta - C'(\bar{q}).$$

By definition of  $\bar{\theta}(\bar{q})$  and  $\underline{\theta}(\bar{q})$ , the first two terms vanish. Thus we obtain  $\int_{\underline{\theta}(\bar{q})}^{\bar{\theta}(\bar{q})} \theta f(\theta)d\theta = C'(\bar{q})$ , where the right-hand side is downward sloping. Moreover, by the nonnegativity of  $\theta f(\theta)$  and that  $[\underline{\theta}(\bar{q}), \bar{\theta}(\bar{q})]$  expands as  $\bar{q}$  increases, the integral is nondecreasing in  $\bar{q}$ , and there exists a constant  $\check{q}$  such that  $\int_{\underline{\theta}(\bar{q})}^{\bar{\theta}(\bar{q})} \theta f(\theta)d\theta = E\theta$  whenever  $\bar{q} \geq \check{q}$ .

Now we discuss the position of the optimal quality limit  $\bar{q}^{FB}$ . When  $\bar{q} = r'(\theta^*)$ ,  $d\Pi^{FB}(\bar{q})/d\bar{q} < 0$ , and therefore the seller tends to increase the quality limit. If there exists at least one quality limit  $\bar{q}$  such that  $d\Pi^{FB}(\bar{q})/d\bar{q} = 0$ , then the optimal level can be obtained by searching over these local maxima; otherwise,  $\bar{q}^{FB} = 0$  is the unique optimal strategy.

If  $C'(r'(\theta^*)) > E\theta$ , then any choice of  $\bar{q}$  will make  $d\Pi^{FB}(\bar{q})/d\bar{q}$  negative, since  $C'(\bar{q}) \geq C'(r'(\theta^*)) > E\theta = \max_{\bar{q}} \int_{\underline{\theta}(\bar{q})}^{\bar{\theta}(\bar{q})} \theta f(\theta)d\theta$ . In this case, any choice of quality limit above  $r'(\theta^*)$  is suboptimal. Because  $\bar{q} \in (0, r'(\theta^*)]$  are all dominated strategies, the optimal choice is  $\bar{q}^{FB} = 0$  if  $C'(r'(\theta^*)) > E\theta$ . When  $C'(\check{q}) > E\theta$ , increasing  $\bar{q}$  after  $\check{q}$  will not change  $\int_{\underline{\theta}(\bar{q})}^{\bar{\theta}(\bar{q})} \theta f(\theta)d\theta$  but will drive up the cost  $C(\bar{q})$ . Hence, any choice above  $\check{q}$  is suboptimal.  $\square$

## Proof of Theorem 2

We shall start with the case when  $R > \theta^*$  and  $\bar{q} > r'(\theta^*)$ . Our strategy is to first ignore the IC and participation conditions for customers outside the interval  $[\underline{\theta}, \tau]$ , i.e., (IC-2), (IC-3), (IC-4), and (PC-2), and then verify that they are satisfied under our proposed menu.

### 1. Proposing the candidate menu

Define  $U(\theta) = \theta q(\theta) - p(\theta) - r(\theta)$ ,  $\forall \theta \in [\underline{\theta}, \tau]$ . Since each served customer should receive at least her reservation utility, we have condition (PC):  $U(\theta) \geq 0$ .

Consider (IC-1) in Eq. (1). For each type  $\theta \in (\underline{\theta}, \tau)$ , the incentive compatibility requires that the payoff is maximized at  $z = \theta$ , and hence the first-order condition yields  $\theta q'(\theta) - p'(\theta) = 0$ ,  $\forall \theta \in (\underline{\theta}, \tau)$ . Differentiating  $U(\theta)$  and plugging in this equality, we have

$$(LO) \quad U'(\theta) = q(\theta) - r'(\theta), \quad \theta \in [\underline{\theta}, \tau].$$

We shall replace constraint (IC-1) in Eq. (1) by (LO), and obtain the necessary conditions for optimality for the modified problem. We will later verify that our proposed schedule satisfies all the imposed constraints and hence it is indeed optimal. Note that  $p(\theta) = \theta q(\theta) - r(\theta) - U(\theta)$ . Observing that the first term in the objective function is independent of the choice of  $(q(\theta), p(\theta))$ ,  $\forall \theta \in [\underline{\theta}, \tau)$ , we can ignore it for the optimization problem. Replacing  $p(\theta)$  by the above expression, the seller's objective becomes  $\max \int_{\underline{\theta}}^{\tau} [\theta q(\theta) - r(\theta) - U(\theta) - c]f(\theta)d\theta$ , subject to (LO) and (PC).

Observe that adding and subtracting constants will not influence the optimal solution, we now remove  $-r(\theta) - c$  from the integrand and add  $-\theta r'(\theta)$  instead. The integrand now becomes  $[\theta q(\theta) - \theta r'(\theta) - U(\theta)]f(\theta)$ , which is equivalent to  $[U(\theta) - \theta U'(\theta)]f(\theta)$  from (LO). Hence, the seller's problem becomes  $\max \int_{\underline{\theta}}^{\tau} [U(\theta) - \theta \frac{dU(\theta)}{d\theta}] f(\theta) d\theta$ , *s.t.*  $U(\theta) \geq 0$ ,  $U(\tau) = 0$ .

We first claim that type- $\underline{\theta}$  customer should not obtain any surplus, i.e.,  $U(\underline{\theta}) = 0$ . Suppose  $\underline{\theta} > 0$ . If  $U(\underline{\theta}) > 0$ , then  $\underline{\theta}q(\underline{\theta}) - p(\underline{\theta}) > r(\underline{\theta})$  and there exists a constant  $\delta$  such that  $\underline{\theta}q(\underline{\theta}) - p(\underline{\theta}) - r(\underline{\theta}) > \delta$ . Consider a customer with type  $\theta$  slightly below  $\underline{\theta}$  such that  $\theta > \underline{\theta} - \frac{\delta}{2\bar{q}}$  and  $r(\theta) > r(\underline{\theta}) - \frac{1}{2}\delta$ . Such a customer exists since  $\underline{\theta} > 0$  and  $r(\cdot)$  is continuous. Now if type- $\theta$  customer chooses  $(q(\underline{\theta}), p(\underline{\theta}))$ , she receives  $(\underline{\theta} - \epsilon)q(\underline{\theta}) - p(\underline{\theta}) > r(\underline{\theta}) - \frac{\delta}{2\bar{q}}q(\underline{\theta}) + \delta > r(\theta) - \frac{1}{2}\delta - \frac{\delta}{2\bar{q}}q(\underline{\theta}) + \delta > r(\theta)$ , because  $q(\underline{\theta}) \leq \bar{q}$ . On the other hand, suppose that  $\underline{\theta} = 0$ . If under the optimal quality-price schedule  $U(\underline{\theta})$  were positive, we can make a uniform shift of prices  $p(\theta)$  while fixing  $q(\theta)$ . This adjustment does not destroy incentive compatibility but strictly increases the seller's profit.

Let  $U(\theta)$  be the state variable, and  $u(\theta) = dU(\theta)/d\theta$  be the control. Through this transformation, the design of the optimal menu of versions can be recast as an optimal control problem and can be solved by use of calculus of variation. The Hamiltonian is given by  $H(\theta) = (-U(\theta) + \theta u(\theta))f(\theta) + \eta(\theta)u(\theta)$ . The adjoint equation is given by  $\frac{d\eta(\theta)}{d\theta} = \frac{-\partial H}{\partial U} = f(\theta)$ , and the transversality condition gives no information. Denote  $\eta(\tau) = e$ , we obtain  $\eta(\theta) = e - F^c(\theta)$ . The necessary condition for optimality is that the Hamiltonian is maximized by the choice of  $u$  since  $H$  is linear in  $u$ .

Consider the coefficient of  $u$  in  $H$ :  $e + \theta f(\theta) - F^c(\theta)$ . If the coefficient of  $u$  were positive, the solution would be unbounded, and hence  $e = 0$  due to the uniqueness of the maximum. Note that  $\theta f(\theta) - F^c(\theta)$  is the derivative of  $-\theta F^c(\theta)$ , and hence from Assumption 1,  $\theta f(\theta) - F^c(\theta) > 0$  if  $\theta > k$ , and  $\theta f(\theta) - F^c(\theta) < 0$  if  $\theta < k$ . The case  $\theta = k$  has measure zero and hence it will not contribute to the objective. If  $\theta f(\theta) - F^c(\theta) > 0$ , there is no maximum since we can take  $u \rightarrow \infty$ . When  $\theta f(\theta) - F^c(\theta) < 0$  we should make  $u$  as negative as possible. But, the boundary conditions  $U(\theta) \geq 0$  on  $[\underline{\theta}, \tau)$  on the other hand they require that  $u(\theta)$  be greater than or equal to zero whenever  $U(\theta) = 0$ . It therefore follows that  $U(\theta) = 0$  for all  $\theta$  in  $[\underline{\theta}, \tau)$  if  $U(\theta) = 0$  is implementable.

Note that this immediately leads to  $q(\theta) = r'(\theta)$  and  $p(\theta) = \theta r'(\theta) - r(\theta)$ ,  $\forall \theta \in [\underline{\theta}, \tau)$ , whenever  $r'(\tau) \leq \bar{q}$ . In this case, a jump in quality occurs at  $\theta = \tau$  and  $\bar{\theta}$  is determined by the minimum of  $R$  and the solution to the equality  $\bar{\theta}\bar{q} - (\tau\bar{q} - r(\tau)) = r(\bar{\theta})$ . From the convexity of  $r(\cdot)$  and that  $r'(\tau) < \bar{q}$ , there exists a unique  $\bar{\theta}$  for any given  $\bar{q}$  and  $\tau$ .

When  $r'(\tau) = \bar{q}$ , the interval  $[\tau, \bar{\theta}]$  degenerates since if a type- $\theta$  customer accepts the version  $(r'(\tau), p(\tau))$  when  $\theta > \tau$ , her payoff will be  $\theta r'(\tau) - (\tau r'(\tau) - r(\tau))$ . But the strict convexity of  $r(\cdot)$  implies that  $r(\theta) > \theta r'(\tau) - (\tau r'(\tau) - r(\tau))$ , and no customer with  $\theta$  higher than  $\tau$  would accept the bundle. This completes the derivation of the proposed menu.

## 2. Checking the necessary and sufficient conditions

Now we check that all other IC and participation conditions are satisfied.

**Checking (IC-1).** Suppose  $\theta \in [\underline{\theta}, \tau)$ . (IC-1) requires that  $\theta r'(z) - (zr'(z) - r(z)) \leq r(\theta)$ ,  $\forall z \in [\underline{\theta}, \tau)$  and  $r(\theta) \geq \theta \bar{q} - (\tau \bar{q} - r(\tau))$ . The former is simply the gradient inequality. For the latter, we have  $r(\tau) - (\tau - \theta)\bar{q} \leq r(\tau) - (\tau - \theta)r'(\tau) \leq r(\theta)$ , where the first inequality follows from that  $\bar{q} \geq r'(\tau)$  and the second one is again the gradient inequality.

**Checking (PC-2) and (IC-2).** We first verify that the participation conditions hold for  $\theta \in [\tau, \bar{\theta}]$ . Recall that by accepting the version  $(\bar{q}, p(\tau))$  both types  $\tau$  and  $\bar{\theta}$  receive their respective reservation utilities, and therefore  $r'(\tau) < \bar{q} < r'(\bar{\theta})$ . If there exists a type  $\theta \in (\tau, \bar{\theta})$  such that  $\theta \bar{q} - (\tau \bar{q} - r(\tau)) < r(\theta)$ , then  $r'(\theta)$  must be greater than  $\bar{q}$ . By monotonicity of  $r(\cdot)$ , the type- $\bar{\theta}$  customer cannot receive  $r(\bar{\theta})$  if she accepts the same version.

Now we consider their incentive compatibility. Given the menu, (IC-2) becomes  $\theta \bar{q} - (\tau \bar{q} - r(\tau)) \geq \theta r'(z) - zr'(z) + r(z)$ ,  $\forall \theta \in [\tau, \bar{\theta}]$ ,  $\forall z \in [\underline{\theta}, \tau]$ . Having established (PC-2), it suffices to show that  $r(\theta) \geq \theta r'(z) - zr'(z) + r(z)$ ,  $\forall \theta \in [\tau, \bar{\theta}]$ ,  $\forall z \in [\underline{\theta}, \tau]$ , which is identical to  $r(\theta) \geq r(z) + (\theta - z)r'(z)$ . Therefore, (IC-2) is true by the convexity of  $r(\cdot)$ .

**Checking (IC-3) and (IC-4).** This follows directly from gradient inequality.

**Checking the necessity of  $p(\tau) = \tau \bar{q} - r(\tau)$ .** Suppose this were not true. Then for type- $\tau$  customer,  $\tau \bar{q} - p(\tau) > r(\tau)$  since her participation condition has to be satisfied. Let  $\delta \equiv \tau \bar{q} - p(\tau) - r(\tau) > 0$ . By continuity and the finiteness of  $\bar{q}$ , there must exist a  $\check{\theta}$  slightly less than  $\tau$  such that  $\check{\theta} > \tau - \frac{\delta}{2\bar{q}}$  and  $r(\check{\theta}) > r(\tau) - \frac{1}{2}\delta$ . The type- $\check{\theta}$  customer is supposed to receive her reservation utility according to the seller's plan. But choosing  $(\bar{q}, p(\tau))$  gives rise to a payoff  $\check{\theta} \bar{q} - p(\tau) = (\tau - \check{\theta})\bar{q} + \tau \bar{q} - p(\tau) = (\tau - \check{\theta})\bar{q} + r(\tau) + \delta > -\frac{1}{2}\delta + r(\tau) + \delta > r(\check{\theta})$ . Thus,  $p(\tau)$  must be  $\tau \bar{q} - r(\tau)$  to avoid profitable deviations.

**Checking the sufficiency.** As the Hamiltonian is linear in  $u$ , it is concave in  $u$  and satisfies the sufficient condition for optimality (Sethi and Thompson [1981, Theorem 2.2]).

## 3. Optimal choice of $\underline{\theta}$ , $\tau$ , and $\bar{\theta}$

We now consider the optimal choice of  $\underline{\theta}$  and  $\tau$ .  $\bar{\theta}$  is determined once we have fixed  $\tau$ .

**Choice of  $\underline{\theta}$ .** Following the proposed quality-price schedule, the seller's net profit from

serving a type- $\theta$  customer is  $p(\theta) - c = \theta r'(\theta) - r(\theta) - c$ , which coincides with  $G(\theta)$ . Since  $G(\theta)$  is positive if and only if  $\theta \geq \theta^*$ , the seller should not sell any version to customers with  $\theta$  below  $\theta^*$ . This suggests that  $\underline{\theta} = \theta^*$ . Furthermore, the choice of  $\underline{\theta}$  will not change either the quality-price schedule for  $\theta \in [\underline{\theta}, \tau]$  or the decision of  $\tau$  and  $\bar{\theta}$ , and thus  $\underline{\theta} = \theta^*$ .

**Choice of  $\tau$ .** Let  $\Xi(\tau)$  denote the profit function of the seller when customers whose types fall into  $[\tau, \bar{\theta}]$  are offered the same version  $(\bar{q}, p(\tau))$  and customers with  $\theta \in [\theta^*, \tau]$  are offered the bundle  $(r'(\theta), \theta r'(\theta) - r(\theta))$ . To indicate the dependence of  $\bar{\theta}$  on  $\tau$ , we shall use  $\bar{\theta} \equiv \bar{\theta}(\tau)$ . Hence  $\Xi(\tau) = (\tau \bar{q} - r(\tau) - c)[F(\bar{\theta}(\tau)) - F(\tau)] + \int_{\theta^*}^{\tau} (\theta r'(\theta) - r(\theta) - c) f(\theta) d\theta$ .

We first consider the case  $\bar{q} \geq \frac{r(R) - r(k)}{R - k}$ . Since  $r(R) \leq R\bar{q} - (k\bar{q} - r(k))$ ,  $\bar{\theta}(\tau)$  is forced to be  $R$  if  $\tau = k$ . Thus,  $\Xi(\tau) = (\tau \bar{q} - r(\tau) - c)[1 - F(\tau)] + \int_{\theta^*}^{\tau} (\theta r'(\theta) - r(\theta) - c) f(\theta) d\theta$ . Using the rule for differentiating under the integral we obtain  $d\Xi(\tau)/d\tau = [\bar{q} - r'(\tau)][1 - F(\tau) - \tau f(\tau)]$ . Since  $\bar{q} > r'(\tau)$  from the definition of  $\tau$ , the sign of  $d\Xi(\tau)/d\tau$  depends only on  $1 - F(\tau) - \tau f(\tau)$ . The point at which the profit function achieves its maximum is independent of the reservation utility. Moreover,  $1 - F(\tau) - \tau f(\tau)$  is the derivative of  $\tau F^c(\tau)$ , which by Assumption 1 has a unique maximum in the interior of  $[0, 1]$ . We conclude that  $\tau = k$  is optimal.

Now we discuss the case  $\bar{q} < \frac{r(R) - r(k)}{R - k}$ . First we assume that  $\bar{\theta}(\tau) \leq R$ , and later we will verify that for the optimality we need not consider other cases. The derivative becomes

$$\frac{d\Xi(\tau)}{d\tau} = [\tau \bar{q} - r(\tau) - c] f(\bar{\theta}(\tau)) \frac{d\bar{\theta}(\tau)}{d\tau} + (\bar{q} - r'(\tau)) [F(\bar{\theta}(\tau)) - F(\tau) - \tau f(\tau)].$$

The term  $\tau \bar{q} - r(\tau) - c$  is nonnegative since  $\tau \bar{q} - r(\tau) - c \geq \tau r'(\tau) - r(\tau) - c > 0$  if  $\tau > \theta^*$ .

To obtain  $\frac{d\bar{\theta}(\tau)}{d\tau}$  we shall fix  $\bar{q}$  and consider two choices  $\tau_1, \tau_2$  of  $\tau$ , and assume that  $\tau_1 < \tau_2$ . The discussion is divided into cases. If  $\bar{\theta}(\tau_1) = R$ , then  $\bar{\theta}(\tau_2) \leq \bar{\theta}(\tau_1)$  as desired. Now assume that  $\bar{\theta}(\tau_2) = R$ . In this case,  $r(R) \leq R\bar{q} - (\tau_2 \bar{q} - r(\tau_2))$ , and hence  $\bar{q} \geq \frac{r(R) - r(\tau_2)}{R - \tau_2}$ . By convexity of  $r(\cdot)$ , we obtain that  $\bar{q} \geq \frac{r(R) - r(\tau_1)}{R - \tau_1}$  as well. Rearranging the above inequality, we conclude that  $\bar{\theta}(\tau_1) = R$  too. Hence in this case,  $\bar{\theta}(\tau_2) = \bar{\theta}(\tau_1)$ .

Finally, let us consider the case when  $\bar{\theta}(\tau_1), \bar{\theta}(\tau_2) \neq R$ . Recall the equality  $r(\bar{\theta}(\tau)) = \bar{\theta}(\tau) \bar{q} - (\tau \bar{q} - r(\tau))$ . From the definition of  $\bar{\theta}$  we have  $r(\bar{\theta}(\tau_1)) = \bar{\theta}(\tau_1) \bar{q} - (\tau_1 \bar{q} - r(\tau_1))$  and  $r(\bar{\theta}(\tau_2)) = \bar{\theta}(\tau_2) \bar{q} - (\tau_2 \bar{q} - r(\tau_2))$ . Let type- $\bar{\theta}(\tau_2)$  customer take the version  $(\bar{q}, \tau_1 \bar{q} - r(\tau_1))$ , i.e. the version designed for customers with  $\theta \in [\tau_1, \bar{\theta}(\tau_1)]$  if  $\tau_1$  is chosen to be the switching customer to accept the same version. The type- $\bar{\theta}(\tau_2)$  customer's payoff becomes  $\bar{\theta}(\tau_2) \bar{q} - (\tau_1 \bar{q} - r(\tau_1)) = r(\bar{\theta}(\tau_2)) + \tau_2 \bar{q} - r(\tau_2) - (\tau_1 \bar{q} - r(\tau_1))$ . By the mean value theorem, there exists a constant  $\tau_3 \in [\tau_1, \tau_2]$  such that  $r(\tau_2) - r(\tau_1) = r'(\tau_3)(\tau_2 - \tau_1)$ . Hence we can rewrite the type- $\bar{\theta}(\tau_2)$  customer's payoff as  $r(\bar{\theta}(\tau_2)) + (\tau_2 - \tau_1)(\bar{q} - r'(\tau_3)) \geq r(\bar{\theta}(\tau_2))$ , where the inequality



follows from that  $\bar{q} \geq r'(\tau_2)$  and the convexity of  $r(\cdot)$ . Thus,  $\bar{\theta}(\tau)$  is decreasing in  $\tau$  and  $\frac{d\bar{\theta}(\tau)}{d\tau} \leq 0$ . We conclude that  $[\tau\bar{q} - r(\tau) - c]f(\bar{\theta}(\tau))\frac{d\bar{\theta}(\tau)}{d\tau} \leq 0$ .

Now consider the second term  $(\bar{q} - r'(\tau))[F(\bar{\theta}(\tau)) - F(\tau) - \tau f(\tau)]$ . Due to  $\bar{q} \geq r'(\tau)$  we only need to consider the sign of the term inside the parentheses. If  $\tau > k$ ,  $F(\bar{\theta}(\tau)) - F(\tau) - \tau f(\tau) \leq 1 - F(\tau) - \tau f(\tau) < 0$ . Thus any  $\tau$  above  $k$  cannot be an optimal solution. This completes the characterization of the optimal menu of contracts when  $R > \theta^*$  and  $\bar{q} > r'(\theta^*)$ . If either condition does not hold, then the seller cannot gather any positive profit.  $\square$

### Proof of Theorem 3

We will follow Jullien [2000] to prove this theorem and hence shall introduce his notation to make a clear connection. Let  $v(\theta, q)$  be the gross utility of type- $\theta$  customer while offered quality  $q$ , and  $s(\theta, q) = v(\theta, q) - c(q)$  be the total surplus from the transaction, and  $w(\theta)$  is the net utility received by the type- $\theta$  customer given the quality-price schedule  $\{q(\theta), p(q(\theta))\}$ . In our model  $v(\theta, q) = \theta q$ ,  $w(\theta) = \theta q(\theta) - p(q(\theta))$ . Since in the production stage  $q > \bar{q}$  is impossible, the production cost is  $c(q) = c$  if  $q \in (0, \bar{q}]$ , and  $c(q) = \infty$  when  $q > \bar{q}$ .

In the sequel, we will verify those relevant conditions required in Jullien [2000], and state and prove the results parallel to Jullien [2000] in our model. The following four lemmas (Lemmas 3-6) are stated and used only in the appendix. The first observation is that the seller never loses money by offering a version:

**Lemma 3.** *Suppose that  $q^*(\theta)$  is an optimal allocation and  $w^*(\theta)$  is the corresponding net utilities. If type- $\theta$  participates, then  $s(\theta, q^*(\theta)) - w^*(\theta) \geq 0$ .*

*Proof.* Let  $T = \{\theta | s(\theta, q^*(\theta)) - w^*(\theta) < 0\}$ . Suppose  $\theta \in T$ . By definitions of  $s(\theta, q^*(\theta))$  and  $w^*(\theta)$ , we have  $\theta q^*(\theta) - c - (\theta q^* - p(q^*)) < 0$ , i.e.,  $p(q^*) - c < 0$ . Therefore, the seller gets a strictly negative profit from the type- $\theta$  customer, for every  $\theta \in T$ .

Suppose now the seller offers instead  $(q^*(\theta), c)$  to every  $\theta \in T$  and keeps every other version the same. For all  $\theta' \notin T$ , choosing version  $(q^*(\theta'), p(q^*(\theta')))$  is still optimal because the prices for  $\theta \in T$  are higher. For  $\theta \in T$ , if the customer chooses any version originally designed for  $\theta' \notin T$ , then the seller gets  $p(q^*(\theta')) - c \geq 0$ ; otherwise, if she chooses a new version  $(q^*(\theta'), c)$ ,  $\theta' \in T$ , by construction the seller would just break even. Thus, in all cases the seller obtains a higher profit, which contradicts the optimality of  $(q^*, p^*)$  if  $T \neq \emptyset$ .  $\square$

We then introduce the following technical definitions.

**Definition 1.**  $(v(\theta, q), r(\theta))$  satisfies Homogeneity (H) if we can find a nondecreasing quality schedule  $\hat{q}(\theta)$  such that  $r'(\theta) = v_\theta(\theta, \hat{q}(\theta))$ ,  $\forall \theta$ .  $(v(\theta, q), r(\theta))$  is a full participation model (FPM) if there exists a tariff  $\hat{c}(q)$  such that  $r(\theta) = \max_q \{v(\theta, q) - \hat{c}(q)\}$ ,  $\forall \theta$ .

We now prove that our model satisfies both conditions given above.

**Lemma 4.** *When  $v(\theta, q) = \theta q$ , and  $r(\theta)$  is differentiable, increasing, and strictly convex, then both (H) and (FPM) are satisfied.*

*Proof.* If we choose  $\hat{q}(\theta) = r'(\theta)$ , then  $r'(\theta) = \hat{q}(\theta) = v_\theta(\theta, \hat{q}(\theta))$  and  $\hat{q}(\theta)$  is increasing by the convexity of  $r(\theta)$ . Hence, (H) is satisfied in our model.

We next show that by setting  $\hat{c}(q) = (r')^{-1}(q)q - r((r')^{-1}(q))$ ,  $r(\theta) = \max_q \{\theta q - \hat{c}(q)\}$ ,  $\forall \theta$ . Recall that  $\hat{q}(\theta) = r'(\theta)$ . Since  $r(\cdot)$  is strictly convex,  $r'(\cdot)$  is strictly increasing and its inverse  $(r')^{-1}(\cdot)$  exists. If we represent  $\hat{c} \equiv \hat{c}(\theta)$  and set  $\hat{c}(\theta) = \theta r'(\theta) - r(\theta)$ , then while checking (IC-1) in the proof of Theorem 2 we know that  $\theta = \operatorname{argmax}_z \{\theta \hat{q}(z) - \hat{c}(z)\}$  and  $r(\theta) = \max_z \{\theta \hat{q}(z) - \hat{c}(z)\}$ . The participation condition corresponds to  $\theta \hat{q}(\theta) - \hat{c}(\theta) = \theta r'(\theta) - [\theta r'(\theta) - r(\theta)] = r(\theta)$ , and hence is satisfied. Therefore  $r(\theta) = \max_z \{\theta \hat{q}(z) - \hat{c}(z)\}$ ,  $\forall \theta$ .

Because there exists a one-to-one correspondence between  $\hat{q}$  and  $\theta$  by the strict monotonicity of  $r'(\cdot)$ , we redefine  $\hat{c}(\theta) = \hat{c}(\hat{q})$ . Note that the existence of  $r''(\cdot)$  gives us the continuity of  $\hat{q}(\cdot)$ . Replacing  $\theta$  by  $(r')^{-1}(\hat{q})$  in  $\hat{c}$  and rename  $\hat{q}$  as the dummy variable  $q$ ,  $\hat{c}(q) = (r')^{-1}(q)q - r((r')^{-1}(q))$  implements  $\{r, \hat{q}\}$ , i.e.,  $r(\theta) = \max_q \{v(\theta, q) - \hat{c}(q)\}$ ,  $\forall \theta$ .  $\square$

Define  $\hat{q}(\theta) = r'(\theta)$  and  $\hat{c}(q) = (r')^{-1}(q)q - r((r')^{-1}(q))$ . Since  $\{\hat{q}, \hat{c}\}$  implements  $r(\theta)$ , we can represent  $\hat{c} \equiv \hat{c}(\theta) = \theta r'(\theta) - r(\theta)$ . Hence in the following we will use  $\hat{c}(q)$  and  $\hat{c}(\theta)$  alternatively for convenience.<sup>14</sup> We now follow Jullien [2000, Section 4]. We slightly modify the original problem in which the seller has the option to serve customers using this technology and charge them at cost  $\hat{c}(\cdot)$ . Note that  $\hat{c}(q)$  has to be equal to  $(r')^{-1}(q)q - r((r')^{-1}(q))$  as defined above. The following lemma describes the set of customers served using the alternate technology  $\hat{c}(\cdot)$ .

**Lemma 5.** *Let the seller be endowed with the two technologies  $c(\cdot), \hat{c}(\cdot)$ . Suppose  $\{q^*(\theta), p^*(\theta)\}$  is an optimal schedule, and  $w^*(\theta)$  is the corresponding customers' utility under this schedule. If  $\hat{c}(q^*(\theta)) \leq c(q^*(\theta))$ , then  $q^*(\theta) = \hat{q}(\theta)$ ,  $p^*(q^*(\theta)) = \hat{c}(\theta)$ , and  $w^*(\theta) = r(\theta)$ . Moreover, only customers with  $\hat{c}(q^*(\theta)) \leq c(q^*(\theta))$  are served by the technology  $\hat{c}(\cdot)$ .*

*Proof.* From Lemma 3,  $s(\theta, q^*(\theta)) = \theta q^*(\theta) - \min\{c(q^*(\theta)), \hat{c}(q^*(\theta))\} \geq w^*(\theta)$ ,  $\forall \theta$ . When  $\hat{c}(q^*(\theta)) \leq c(q^*(\theta))$ , we obtain  $0 \leq \theta q^*(\theta) - \min\{c(q^*(\theta)), \hat{c}(q^*(\theta))\} - w^*(\theta) \leq \theta q^*(\theta) - \hat{c}(q^*(\theta)) - w^*(\theta) \leq r(\theta) - w^*(\theta)$ , where the last inequality follows from that  $r(\theta) = \max_q \{\theta q - \hat{c}(q)\}$ . Recall that a customer should receive at least her reservation utility, i.e.,  $w^*(\theta) \geq r(\theta)$ . Therefore,  $w^*(\theta) = r(\theta)$ .

Now we show that when  $\hat{c}(q^*(\theta)) \leq c(q^*(\theta))$ ,  $q^*(\theta) = \hat{q}(\theta)$ . We only need to consider the case where  $\hat{c}(q^*(\theta)) \leq c(q^*(\theta))$  occurs in an open interval: if this occurs only at isolated points, then the contribution of these customers to the seller's profit is negligible. Assume

---

<sup>14</sup>Note also that for any optimal schedule  $\{q(\theta), p(q(\theta))\}$ , we can also represent  $p(q(\theta))$  by  $p(\theta)$  according to the revelation principle. This convention is also adopted here.

$q^*(\theta) > \hat{q}(\theta)$  for some  $\theta$  and define  $\delta = q^*(\theta) - \hat{q}(\theta) > 0$ . Since  $\hat{c}(q^*(\theta)) \leq c(q^*(\theta))$ , we know that  $w^*(\cdot) = r(\cdot)$  in a neighborhood of  $\theta$ . Thus, there must exist a  $\theta_1 > \theta$  that satisfies  $w^*(\theta_1) = r(\theta_1)$  and  $\hat{q}(\theta_1) < \hat{q}(\theta) + \frac{1}{2}\delta$ . By choosing type- $\theta$ 's version, the type- $\theta_1$  customer receives  $\theta_1 q^*(\theta) - p^*(q^*(\theta)) = (\theta_1 - \theta)q^*(\theta) + r(\theta) > (\theta_1 - \theta)\hat{q}(\theta_1) + r(\theta) \geq r(\theta_1)$ , where the strict inequality is by construction, and the last inequality follows from that  $\hat{q} \equiv r'$  and the convexity of  $r(\cdot)$ . This violates the IC condition for  $\theta_1$ . Similarly, we can show that  $q^*(\theta) < \hat{q}(\theta)$  is also impossible. Since  $r(\theta) = w^*(\theta) = \theta\hat{q}(\theta) - p^*(\hat{q}(\theta)) = \theta r'(\theta) - p^*(\theta)$ , we have  $p^*(\hat{q}(\theta)) = \theta r'(\theta) - r(\theta)$ , identical to  $\hat{c}$ . Thus,  $p^*(\theta) = \hat{c}(\theta)$  if  $\hat{c}(q^*(\theta)) \leq c(q^*(\theta))$ . When  $\hat{c}(q^*(\theta)) > c(q^*(\theta))$ , it is optimal to use the original technology  $c(\cdot)$ , and hence only customers with condition  $\hat{c}(q^*(\theta)) \leq c(q^*(\theta))$  are served by the alternate technology.  $\square$

Finally, we claim that any optimal schedule must include an interval of customers.

**Lemma 6.** *If  $\{q^*(\theta), p^*(\theta)\}$  is optimal, it must serve an interval of customers.*

*Proof.* We first prove that  $\hat{c}(q)$  is strictly increasing. Let us recall the definition of  $\hat{q}(\theta) = r'(\theta)$  and  $\hat{c}(\cdot)$  can be represented as a function of  $\theta$ :  $\hat{c}(\theta) = \theta r'(\theta) - r(\theta)$ . Differentiating  $\hat{c}$  by  $\theta$ , we have  $\hat{c}'(\theta) = \theta r''(\theta)$ , which is positive when  $\theta > 0$  by the strict convexity of  $r(\theta)$ . Therefore,  $\hat{c}$  is strictly increasing in  $\theta$ . The result follows from the strict monotonicity of  $\hat{q}(\theta) = r'(\theta)$ .

Following Jullien [2000, Section 4], we can always assume full participation with the alternate tariff  $\hat{c}$ , in which  $q^*(\theta)$  must be monotonic. If a customer  $\theta$  is not included, then we can have the seller offer  $(\hat{q}(\theta), \hat{c}(\theta))$  to her, and the incentives of both parties are verified. Moreover, the payoff is equivalent under such a modification.

We can apply Lemma 5 to characterize the set of excluded customers. Note that in our model  $c(q) = c$  if  $q \in (0, \bar{q}]$ , and  $c(q) = \infty$  when  $q > \bar{q}$ . Define  $J = \{\theta | q^*(\theta) \leq \bar{q}\}$  as the set of customers that receive a version with quality less than  $\bar{q}$ . By monotonicity of  $q^*$   $J$  should be an interval  $[0, \beta]$  if it does not degenerate. Since  $\hat{c}$  and  $q^*$  are both monotonic, the set  $J_E = \{\theta : \hat{c}(q^*(\theta)) \leq c, \theta \in J\}$  is either empty or an interval  $[0, \alpha]$  where  $\alpha \leq \beta$ . According to Lemma 5, in  $J$  the seller can at most exclude an interval of customers that starts from  $\theta = 0$  ( $[0, \alpha]$  as labelled); otherwise, she has to serve all customers in  $J$  under the schedule  $q^*$ . When  $\theta \notin J$ ,  $q^*(\theta) > \bar{q}$ , in which case  $\hat{c}(q^*(\theta)) < c(q^*(\theta)) = \infty$ , and hence by Lemma 5 the customers not in  $J$  should be excluded. Note that from the monotonicity of  $q^*(\theta)$ , this set is either empty or an interval  $(\beta, R]$ . Combining all above, the exclusion can either be an interval  $[0, \alpha]$  with  $\alpha \leq \beta$ , or  $(\beta, R]$ , and therefore at optimality no intermediate exclusion is considered, i.e., the seller must serve an interval of customers.  $\square$

Since the optimal schedule must serve an interval of customers, the schedule proposed in Theorem 2 remains optimal when the seller is allowed to exclude customers arbitrarily.  $\square$

### Proof of Theorem 4

Recall that  $\underline{\theta}^{SB} = \theta^*$ , and the boundary points for the first-degree price discrimination are the two roots of  $\theta\bar{q} - r(\theta) - c = 0$ . Plugging  $\theta = \theta^*$  in this equation, we have  $\theta^*\bar{q} -$

$r(\theta^*) - c \geq \theta^* r'(\theta^*) - c = 0$ , and hence  $\underline{\theta}^{FB} \leq \underline{\theta}^{SB}$ . For the right boundary points, we have  $\bar{\theta}^{SB} \bar{q} - \tau \bar{q} + r(\tau) \geq r(\bar{\theta}^{SB})$ , and the equality holds when  $R$  has not been hit. Thus,  $\bar{\theta}^{SB} \bar{q} - r(\bar{\theta}^{SB}) - c \geq \tau \bar{q} - r(\tau) - c \geq \tau r'(\tau) - r(\tau) - c \geq 0$ , where the second inequality follows from the choice of  $\tau$ , and the third inequality is because  $\tau \geq \underline{\theta}^{SB} = \theta^*$ . When  $\bar{\theta}^{SB} = R$ , we obtain  $R \bar{q} - r(R) - c \geq \tau \bar{q} - r(\tau) - c \geq \tau r'(\tau) - r(\tau) - c \geq 0$ , and therefore  $\bar{\theta}^{FB} = R$  as well. Hence, we also have  $\bar{\theta}^{SB} \leq \bar{\theta}^{FB}$ . The last two observations follow directly from the optimal schedules characterized in Proposition 1 and Theorem 2.  $\square$

### Proof of Theorem 5

Consider the first-degree price discrimination. While replacing  $\bar{q}_1$  by  $\bar{q}_2$ , the difference between the total surplus and the reservation utility becomes larger. Therefore, the seller collects more profit from every served customer. Since the set of served customers is strictly larger, the total profit in the production stage can only go higher.

Now we switch to the second-degree price discrimination. Let  $\tau(\bar{q}_1)$  be the optimal starting point of offering the version of the highest quality when the quality limit is  $\bar{q}_1$  and assume the seller chooses the same  $\tau(\bar{q}_1)$  under quality limit  $\bar{q}_2$ . First, because the choice of  $\tau$  affects neither the schedule offered before this switching point nor the cutoff point of discarding the low-end customers, the seller gets exactly the same profits from every customer with  $\theta \leq \tau(\bar{q}_1)$ . Note that when  $\theta \in [\theta^*, \tau(\bar{q}_1))$  the incentive compatibility and participation constraints require that  $q(\theta) = r'(\theta)$  and exactly the same price schedule.

Regarding the set of customers that accept a common version, she gains in two aspects. First, the price of this common version is strictly higher because  $\tau(\bar{q}_1) \bar{q}_2 - r(\tau(\bar{q}_1)) > \tau(\bar{q}_1) \bar{q}_1 - r(\tau(\bar{q}_1))$ . Second, more high-end customers are willing to purchase this version compared to the case with  $\bar{q}_1$ . Define  $\bar{\theta} \equiv \bar{\theta}(\tau(\bar{q}_1))$  for ease of notation. If  $\bar{\theta}$  has hit the boundary  $R$ , then we have  $r(R) \leq R \bar{q}_1 - [\tau(\bar{q}_1) \bar{q}_1 - r(\tau(\bar{q}_1))]$ , which gives us  $\bar{q}_1 \geq \frac{r(R) - r(\tau(\bar{q}_1))}{R - \tau(\bar{q}_1)}$ . Now if we replace  $\bar{q}_1$  by  $\bar{q}_2$ , we obtain  $\bar{q}_2 > \bar{q}_1 \geq \frac{r(R) - r(\tau(\bar{q}_1))}{R - \tau(\bar{q}_1)} \Rightarrow r(R) < R \bar{q}_2 - [\tau(\bar{q}_1) \bar{q}_2 - r(\tau(\bar{q}_1))]$ , which implies that with  $\bar{q}_2$  no high-end customers is excluded as well. The only case left here is  $\bar{\theta} < R$ . In this case,  $r(\bar{\theta}) = \bar{\theta} \bar{q}_1 - [\tau(\bar{q}_1) \bar{q}_1 - r(\tau(\bar{q}_1))]$ , and hence  $\frac{r(\bar{\theta}) - r(\tau(\bar{q}_1))}{\bar{\theta} - \tau(\bar{q}_1)} = \bar{q}_1 < \bar{q}_2$ . This inequality implies that customers with  $\theta \in [\tau(\bar{q}_1), \bar{\theta}]$  are served with  $\bar{q}_2$  in place of  $\bar{q}_1$ . Therefore, the seller serves more high-end customers under  $\bar{q}_2$ .

By adopting the same starting point  $\tau(\bar{q}_1)$ , the seller gains more with  $\bar{q}_2$  in place of  $\bar{q}_1$ . If she chooses the starting point optimally, her profit can only be higher.  $\square$

### Proof of Theorem 6

Let  $S(\theta) = \arg \max_z \{\theta q(z) - p(z)\} = \theta q(\theta) - p(\theta)$  denote the utility of type- $\theta$  customer, where the second equality follows from incentive compatibility. Define  $U(\theta) \equiv S(\theta) - r(\theta) \geq 0$ , where the inequality follows from the participation constraint. Incentive compatibility requires that  $q(\theta)$  be monotonic, and local optimality implies that  $\frac{dU(\theta)}{d\theta} = q(\theta) - r'(\theta)$ . Recall that  $p(\theta) = \theta q(\theta) - S(\theta) = \theta q(\theta) - U(\theta) - r(\theta)$ . For ease of illustration, we first assume that the seller serves an interval of customers  $[\theta_0, \theta_1]$ .

The seller's optimization problem is

$$\max_{q(\theta), U(\theta)} \int_{\theta_0}^{\theta_1} \{\theta q(\theta) - U(\theta) - r(\theta) - c(q(\theta))\} f(\theta) d\theta,$$

s.t.  $dU(\theta)/d\theta = q(\theta) - r'(\theta)$ , and  $U(\theta) \geq 0, \forall \theta \in [\theta_0, \theta_1]$ .  $q$  is the control variable and  $U$  is the state variable. The Hamiltonian is defined as  $H(U, q, \mu, \theta) = \{\theta q(\theta) - U(\theta) - r(\theta) - c(q(\theta))\} f(\theta) + \mu [q(\theta) - r'(\theta)]$ , where  $\mu$  is the costate variable, and the Lagrangian is  $L = H(U, q, \mu, \theta) + \delta(\theta)U(\theta)$ . The optimal solution should jointly satisfy

$$\begin{aligned} \frac{\partial H}{\partial q} &= (\theta - c'(q))f(\theta) + \mu(\theta) \geq 0, \\ \frac{d\mu}{d\theta} &= -\frac{\partial L}{\partial U} = f(\theta) - \delta(\theta), \\ \frac{dU(\theta)}{d\theta} &= q(\theta) - r'(\theta), \\ \delta(\theta)U(\theta) &= 0, \delta(\theta) \geq 0, U(\theta) \geq 0, \\ \mu(\theta_0)U(\theta_0) &= 0, \mu(\theta_0) \leq 0, \mu(\theta_1)U(\theta_1) = 0, \mu(\theta_1) \geq 0, \end{aligned}$$

where the first inequality follows from first-order condition, the second is the costate equation, the third is the local optimality, the fourth is the complementary slackness, and the last set of equations are transversality conditions.

We first argue that participation constraints cannot be binding within an interval, say,  $[\theta_2, \theta_3]$ . If this were the case,  $q(\theta) = r'(\theta), \forall \theta \in [\theta_2, \theta_3]$ . However, incentive compatibility requires  $q(\theta)$  be increasing, i.e.,  $r'(\theta) \geq 0$  in the interior of  $[\theta_2, \theta_3]$ . If  $r(\theta)$  is concave, this is impossible. If  $r(\theta) = a\theta$ , where  $a \geq 0$ , then it must be that  $q(\theta) = a, p(\theta) = 0, \forall \theta \in [\theta_2, \theta_3]$ . By doing so the seller would lose money due to the positive variable cost  $c(q(\theta))$ . Thus, The full rent extraction region degenerates. This argument continues to hold when we allow arbitrary exclusion. Therefore, we conclude that when  $r(\theta)$  is linear or concave, participation constraints cannot be binding within an interval.

Therefore, if versioning occurs at optimality, it must come from the first-order condition of Hamiltonian. If  $c(q(\theta)) = c$ ,  $q(\theta)$  does not get into  $\partial H/\partial q$ . Therefore, the solution has to be bang-bang: when  $\theta f(\theta) + \mu(\theta) < 0$ , type- $\theta$  customer should be excluded; otherwise, she should be offered the quality limit. Versioning cannot appear in this case.

Now suppose  $c(q)$  is linear or concave. When participation constraint is not binding,  $\delta(\theta) = 0$  from complementary slackness. Therefore, the costate equality implies that the costate equation  $\frac{d\mu}{d\theta} = f(\theta) \Leftrightarrow \mu(\theta) = F(\theta) - A$ , where  $A$  is a constant. This interval could either start with 0 or end with 1 (from Maggi and Rodriguez-Clare [1995]). In these two cases, the transversality condition yields respectively  $\mu(\theta) = F(\theta)$  and  $\mu(\theta) = F(\theta) - 1$ . Plugging in these two values in the binding first-order condition, we have either  $c'(q(\theta)) = \theta + \frac{F(\theta)}{f(\theta)}$  or  $c'(q(\theta)) = \theta - \frac{1-F(\theta)}{f(\theta)}$ . When the regularity condition holds, the right-hand sides are both

increasing. Thus, if  $q(\theta)$  is increasing (from the local optimality condition), it must be that  $c(q)$  is convex. This contradicts our assumption on  $c(q)$ . The proof is now complete.  $\square$

## References

- H. K. Bhargava and V. Choudhary. Information goods and vertical differentiation. *Journal of Management Information Systems*, 18:89–96, 2001.
- H. K. Bhargava and V. Choudhary. Economics of an electronic intermediary with aggregation benefits. *Information Systems Research*, 18:22–36, 2004.
- A. Ghose and A. Sundararajan. Software versioning and quality degradation? an exploratory study of the evidence. Working Paper CeDER-05-20, Center for Digital Economy Research, New York University, 2005.
- K.-W. Huang and A. Sundararajan. Pricing models for on-demand computing. Working paper, New York University, 2005.
- B. Jing. Market segmentation for information goods with network externality. *Economic Letters*, forthcoming, 2006.
- R. Jones and H. Mendelson. Product and price competition for information goods. Working paper, Stanford University, 1998.
- B. Jullien. Participation constraints in adverse selection models. *Journal of Economic Theory*, 93:1–47, 2000.
- J. Juran and F. Gryna. *Juran's Quality Control Handbook*. McGraw-Hill, USA, 1951.
- M. Lariviere. A note on probability distributions with increasing generalized failure rates. *Operations Research*, 54:602–604, 2006.
- G. Maggi and A. Rodriguez-Clare. On countervailing incentives. *Journal of Economic Theory*, 66:238–263, 1995.
- E. Maskin and J. Riley. Monopoly with incomplete information. *RAND Journal of Economics*, 15(2):171–196, 1984.
- M. Mussa and S. Rosen. Monopoly and product quality. *Journal of Economic Theory*, 18:301–317, 1978.
- S. Raghunathan. Software editions: An application of segmentation theory to the packaged software market. *Journal of Management Information Systems*, 17:87–114, 2000.
- F. Riggins. Market segmentation and information development costs in a two-tiered fee-based and sponsorship-based web site. *Journal of Management Information Systems*, 19:69–86, 2002.
- S. Sethi and G. L. Thompson. *Optimal control theory: Applications to management science*. Martinus Nijhoff Publishing, Boston, 1981.
- C. Shapiro and H. Varian. *Information Rules*. Harvard Business School Press, NY, 1998.
- A. Sundararajan. Managing digital piracy: pricing and protection. *Information Systems Research*, 15:287–304, 2004.
- T. Weber. Mixed differentiation of information goods under incomplete information. *Proceedings of the 23rd International Conference on Information Systems*, L. Applegate, R. D. Galliers, and J. I. DeGross (Eds.), Barcelona, pages 81–94, 2002.