Dynamic Price Discrimination and Quality Provision Based on Purchase History

Sun, Ching-jen

Deakin University

October 2007

Dynamic Price Discrimination and Quality Provision Based on Purchase History

Ching-jen Sun*

August 06, 2008

Abstract

This paper develops a general two-period model of product line pricing with customer recognition. Specifically, we consider a monopolist who can sell vertically differentiated products over two periods to heterogeneous consumers. Each consumer demands one unit of the product in each period. In the second period, the monopolist can condition the price-quality offers on the observed purchasing behavior in the first period. In this setup, the monopolist can price discriminate consumers not only by quality, but also by purchase history. Several interesting results are derived. First, we fully characterize the monopolist's optimal pricing strategy when there are two types of consumers, and a simple condition is given to determine whether the monopolist will price discriminate by quality in the first period. We compare it to the case when there is no customer recognition or the firm is able to commit to its future actions. When the type space is a continuum, we show that there is no fully separating equilibrium, and some properties of the optimal contracts (price-quality pairs) are characterized within the class of partitional PBE.

JEL classification: D42; L11

Keywords: Price discrimination; Supermodularity; Submodularity; Behavior-Based Pricing; Ratchet Effect; Bunching.

1 Introduction

Thanks to rapid progress in information technology, online companies, banks, airlines, and grocery stores commonly collect individual information, track consumers' purchase histories, and use this information to identify consumers and offer different prices and personalized products to them accordingly (Taylor (2004) and Turow et al. (2005)). Amazon.com, for example, offers the same DVDs to different customers at different prices based on their purchase histories. One customer deleted the cookies on his computer that identified him as a regular customer, and the price of a DVD fell from \$26.24 to \$22.74 (Streitfeld 2000). To fine-tune its marketing strategies, Harrah's Entertainment Inc., one of the top U.S. casino operators, mines data on its gamblers from

*Department of Economics, The Ohio State University, 464 Arps Hall, 1945 North High Street, Columbus, OH 43210. Tel.: (614)292-5584, e-mail: sun.161@osu.edu. I am grateful to James Peck for his guidance. Also I would like to acknowledge Jan Bouckaert, Huanxing Yang, Lixin Ye and participants at the fifth International IO conference for their helpful comments. All remaining errors are my own.

Total Rewards cards that customers insert into slot machines before they play (Binkley (2000)).[1] Through the personalization and content management software implemented in its website AA.com, American Airlines is able to analyze customer profiles and provide customized offers (Turow et al. (2005)). Victoria's Secret mails catalogues that provide different discounts to different groups of consumers (Weiss and Mehrotra (2001)). A recent article from the Center of Consumer Reports shows that major travel sites adjust their fares in a very sophisticated fashion.[2]

In this paper, we examine "behavior-based price discrimination" (coined by Fudenberg and Villas-Boas (2006)) by studying a dynamic price discrimination model with product design for nondurable goods. Specifically, we consider a monopolist who can sell vertically differentiated products over two periods to heterogeneous consumers. Each consumer demands one unit of the product in each period. In the first period, the monopolist offers a menu of contracts (price-quality pairs), and consumers make their first-period purchase decisions. In the second period, consumers are partitioned into different segments based on their first-period choices, and the monopolist is allowed to offer different contracts to different segments. In this way, the monopolist can discriminate consumers by offering multiple qualities at different prices within each period or by tracking their purchase history and offering different menus of contracts conditional on consumers' previous purchase behavior. Among other questions, we ask when the monopolist will or will not price discriminate consumers either by quality or purchase history and what the quality provision dynamics looks like in equilibrium.

Several interesting results are derived in both discrete and continuous models. First we fully characterize the optimal pricing strategy employed by the monopolist when there are two types of consumers, then we give a simple condition to determine whether the monopolist will price discriminate consumers by quality in the first period. The nature of the equilibria varies with the patience of consumers and the structure of the social surplus function. When the social surplus function is log submodular or consumers and the firm are patient, the firm will forgo the opportunity to price discriminate by quality and offer only one quality in the first period.[3] If the social surplus function is log supermodular, consumers are impatient and the fraction of high types is in between, the firm will offer the product with multiple qualities in the first period. Moreover, the firm will price discriminate consumers by their purchase history if the fraction of high types is high or the social surplus function is log supermodular, the fraction of high types is in between and consumers are impatient.

When the type space is a continuum, we show that, due to the well-known *"ratchet effect"*, there exists no fully separating equilibrium, even though full separation is feasible and could be profitable in the static model. Some properties of the optimal menu of contracts are derived within the class of partitional equilibria. First, we provide an upper bound on the number of varieties of the product in the first period. Hence, in equilibrium the variety of the product is way below the types of consumers. Analogous to the model with two consumer types, the firm will offer only one quality in the first period if the social surplus function is log submodular or the firm and consumers are patient. Furthermore, if it is optimal for the firm to offer only one quality in the first period, the optimal market coverage in the first period is smaller than that in the static model. In equilibrium

[1] The Total Rewards card collects gamblers' characteristics like age and gender and playing habits like how much they spend per trip, what their favorite games are, and even how fast they pull a slot-machine lever.

[2] See http://www.consumerwebwatch.org/dynamic/e-commerce-investigation-dynamic-pricing.cfm

[3] A function $f(x, y) \in C^2$ is said to be log submodular if $f_{xy}f - f_x f_y < 0$.

there are some high-type consumers choosing to *downgrade* the product in the second period, a phenomenon that has never been addressed in the literature. We show that the first-period quality or market coverage is not necessarily monotone with the discount factor δ.

Our research contributes to two strands of the literature on price discrimination: the literature on static product line pricing and the literature on behavior-based price discrimination. This paper is the first study to examine monopoly pricing and product line design dynamics with customer recognition. It differs from the models of static product line design by looking at the dynamics of quality provision and allowing the firm to track consumers' purchase histories and offer customized products. In this regard, we extend the general model of Anderson and Dana (2006) to a dynamic environment. From the viewpoint of product line design, our study extends the single-product models of behavior-based price discrimination by allowing the firm to offer multiple qualities. Furthermore, our model applies to more general environments than those studied in the behavior-based price discrimination literature.

A variety of models can be categorized to the literature on product line pricing.[4] Mussa and Rosen (1978) study the optimal product line design problem for a monopolist and obtain two important observations. Their first observation is that optimal quality provision is distorted downward due to the information rent effect. Their second observation is that without a quality constraint, it is optimal for the monopolist to offer multiple qualities under mild conditions. Stokey (1979) asks whether a monopolist who can commit to future prices and faces a cohort of consumers with heterogeneous tastes might price discriminate consumers by cutting the price over time. She shows that generically the monopolist's optimal strategy is to forego the opportunity to price discriminate and commit to a fixed price over time. If we treat the goods delivered at different dates as the product with different qualities, Stokey's model of intertemporal price discrimination can be transformed into a standard product line pricing problem. Hence, Stokey shows that, in contrast to Mussa and Rosen (1978), it is optimal for the monopolist to offer just one quality.[5]

Anderson and Dana (2006) develop a general model integrating the seemingly unrelated models of product line pricing and derive an elegant condition, called the increasing percentage differences condition or log supermodularity,[6] to determine whether price discrimination is optimal. More specifically, they show that price discrimination by offering multiple qualities is profitable if and only if the percentage change in social surplus from product upgrades is increasing in consumers' willingness to pay.

Behavior-based price discrimination (BBPD) has attracted more attention in the literature recently.[7] Most papers restrict their analyses to the case of a single quality. Hart and Tirole (1988) and Villas-Boas (2004) study price dynamics for a monopoly. Acquisti and Varian (2005) and Taylor (2004) address the issue of consumer privacy. Some authors study dynamic price competition with

[4]The examples include Mussa and Rosen (1978), Johnson and Myatt (2003) and Villas-Boas (1998) on product design decisions, Stokey (1979) on intertemporal price discrimination with commitment, Deneckere and McAfee (1996) and McAfee (2006) on damaged goods, Courty and Li (2000) and Gale and Holmes (1992) on advance purchase discounts, Bhargava and Choudhary (2001) and Varian (2001) on versioning information goods, and Anderson and Song (2004) and Nevo and Wolfram (2002) on coupons.

[5]By integrating the models in Mussa and Rosen (1978) and Stokey (1979) into a common framework, Salant (1989) provides an explanation for these two inconsistent findings.

[6]See Johnson and Myatt (2003) and McAfee (2006) also.

[7]See Fudenberg and Villas-Boas (2006), Armstrong (2006) and Stole (2004) for the recent developments on this topic.

customer recognition (Chen (1997), Fudenberg and Tirole (2000), Taylor (2003), and Villas-Boas (1999)). Only a few papers examine BBPD with multiple products. Fudenberg and Tirole (1998) study a monopoly model of upgrades and buy-backs for a durable good. Zhang (2005) explores firms' endogenous product design in a two-period poaching model. Both models restrict their analyses to either linear or quadratic utility functions and assume that firms can only offer one product in the first period. In contrast, our model allows the monopolist to offer multiple products in either period, and the analysis applies to general utility functions with a mild condition.

The rest of the paper is organized as follows. In Section 2, we study the model with two types of consumers. Section 3 analyzes the situation when the type space is a continuum. Section 4 concludes.

2 The Model with Two Consumer Types

2.1 The Static Model with Exogenously Given Qualities

As a preliminary, we solve for the monopolist's optimal static (one-period) pricing strategy. Consider a monopolist who can sell either or both of two products, one with exogenously given quality $\underline{q}$ and another with exogenously given quality $\overline{q}$. The marginal costs are $c(\underline{q})$ and $c(\overline{q})$ respectively. Let $p(q)$ denote the price of the product with quality q, where $q = \overline{q}$ or $\underline{q}$. There are two consumer types, θ_H and θ_L, and each consumer has unit demand. Let α indicate the fraction of high types θ_H in the population. The monopolist acts to maximize profit, and consumers act to maximize their consumer surplus, $u(q, \theta_H) - p(q)$ and $u(q, \theta_L) - p(q)$ respectively. We adopt the following standard assumption:

Assumption 1 *(i)* $u(\overline{q}, \theta_H) > u(\overline{q}, \theta_L) > c(\overline{q})$, *(ii)* $u(\underline{q}, \theta_H) > u(\underline{q}, \theta_L) > c(\underline{q})$ *(iii)* $u(\overline{q}, \theta_H) - u(\overline{q}, \theta_L) > u(\underline{q}, \theta_H) - u(\underline{q}, \theta_L)$, *(iv)* $u(\overline{q}, \theta_L) - u(\underline{q}, \theta_L) > c(\overline{q}) - c(\underline{q})$.

(i) and (ii) assume that the high-type consumers value the product more than the low-type consumers. (iii) is the single-crossing property, and (iv) implies that if the monopolist were serving only the low types, then it would choose to sell them the high quality product.

For this one-period maximization problem, the firm can choose to offer either one or two qualities of the product to consumers. If the firm chooses to offer two qualities to consumers, then from the single crossing property, it must sell $\overline{q}$ to the high-type consumers and $\underline{q}$ to the low-type consumers. Moreover, the firm extracts all consumer surplus from the low-type consumers ($p(\underline{q}) = u(\underline{q}, \theta_L)$); otherwise the profit is not maximized. To avoid the deviation of the high-type consumers from taking $\underline{q}$, $p(\overline{q})$ can be at most equal to $u(\underline{q}, \theta_L) + u(\overline{q}, \theta_H) - u(\underline{q}, \theta_H)$. Thus the highest profit from offering two qualities is

$$\pi = \alpha[u(\underline{q}, \theta_L) + u(\overline{q}, \theta_H) - u(\underline{q}, \theta_H) - c(\overline{q})] + (1 - \alpha)[u(\underline{q}, \theta_L) - c(\underline{q})].$$

If the firm chooses to just offer one quality, then he will offer $\overline{q}$, the more efficient one according to our assumption. In this case, the firm can decide to either charge a high price to attract only the high-type consumers to purchase or charge a lower price to cover the whole market. Selling to high-type consumers only, the highest profit the firm can get is $\pi = \alpha[u(\overline{q}, \theta_H) - c(\overline{q})]$. On the other hand, by charging a price $p(\overline{q}) = u(\overline{q}, \theta_L)$ the firm can sell to the whole market and get

$\pi = u(\overline{q}, \theta_L) - c(\overline{q})$. The magnitude of α determines which of these three strategies is optimal. Define $S(q, \theta) = u(q, \theta) - c(q)$, the surplus from selling a product of quality q to type θ, and

$$\begin{aligned} \Delta &= \frac{S(\overline{q},\ \theta_L) - S(\underline{q},\ \theta_L)}{S(\overline{q},\ \theta_H) - S(\underline{q},\ \theta_H)} < 1 \\ \Lambda(q) &= \frac{S(q,\ \theta_L)}{S(q,\ \theta_H)} < 1, \end{aligned}$$

where $\Lambda(q)$ is the ratio of the low type's total surplus to the high type's total surplus, and Δ is the ratio of the surplus increases from product upgrades. A straightforward comparison shows that $\Delta > \Lambda(\overline{q})$ if $\Lambda(\overline{q}) > \Lambda(\underline{q})$, and $\Delta \leq \Lambda(\overline{q})$ if $\Lambda(\overline{q}) \leq \Lambda(\underline{q})$. The following proposition characterizes the monopolist's optimal static pricing strategy.

Proposition 1 *In the static (one-period) model with two consumer types, the monopolist's optimal pricing strategy is characterized as follows:*

- *Case 1.* $\Lambda(\overline{q}) > \Lambda(\underline{q})$. *The monopolist will offer only the high-quality product* $\overline{q}$, *at the price* $p = u(\overline{q}, \theta_H)$ *if* $\alpha \geq \Lambda(\overline{q})$, *and* $p = u(\overline{q}, \theta_L)$ *otherwise.*

- *Case 2.* $\Lambda(\overline{q}) \leq \Lambda(\underline{q})$. *In this case,* $\Delta \leq \Lambda(\overline{q}) \leq \Lambda(\underline{q})$ *and the monopolist will offer (i) only* $\overline{q}$ *at* $p = u(\overline{q}, \theta_H)$ *if* $\alpha \geq \Lambda(\underline{q})$, *(ii) only* $\overline{q}$ *at* $p = u(\overline{q}, \theta_L)$ *if* $\alpha \leq \Delta$, *and (iii) both qualities with* $p(\underline{q}) = u(\underline{q}, \theta_L)$ *and* $p(\overline{q}) = u(\underline{q}, \theta_L) + u(\overline{q}, \theta_H) - u(\underline{q}, \theta_H)$ *if* $\Delta \leq \alpha \leq \Lambda(\underline{q})$.

This result is intuitive. When the fraction of high types in the population is high, it is the monopolist's best strategy to simply charge a high price, $p = u(\overline{q}, \theta_H)$, and sell to the high-type consumers only. On the other hand, if the fraction of low types is high, then it is more profitable to offer a low price and cover the whole market. The firm will price discriminate consumers by offering multiple qualities only if the so-called increasing percentage differences condition $\frac{S(\overline{q},\ \theta_L) - S(\underline{q},\ \theta_L)}{S(\overline{q},\ \theta_H) - S(\underline{q},\ \theta_H)} \leq \frac{S(\underline{q},\ \theta_L)}{S(\underline{q},\ \theta_H)}$ or $\frac{S(\overline{q},\ \theta_L)}{S(\underline{q},\ \theta_L)} \leq \frac{S(\overline{q},\ \theta_H)}{S(\underline{q},\ \theta_H)}$ holds (Case 2). This condition says that the ratio of the high type's total surplus to the low type's total surplus is increasing in quality or equivalently that the percentage change in social surplus from product upgrades is increasing in the consumer type.[8]

2.2 Dynamic Optimization with Customer Recognition

We now study the dynamic version of the basic model. We will assume in the rest of the paper that the firm cannot commit itself to second-period behavior. Suppose now that the firm can sell either one or both products over two periods and that consumers have unit demand in each period. The firm acts to maximize the discounted value of its profits using a discount factor $\delta \in (0, 1)$. In the first period, the firm chooses the prices and qualities offered in the market, then consumers make purchase decisions. In the second period, the firm offers prices and qualities that can be conditioned on consumers' purchase histories. Consumer θ's behavior is to maximize the discounted sum of per-period utilities $u_t(q, \theta)$, where for a given period-t payment p_t, $u_t(q, \theta) = u(q, \theta) - p_t$. Consumers use the same discount factor δ as the monopolist.

[8] See Anderson and Dana (2006) for a graphical explanation.

We solve for a PBE of this game. In the second period, after observing the purchase history in the first period, the firm chooses the second-period menus optimally given its beliefs about consumer types in different segments. The beliefs should be consistent with the consumers' first-period choices. Similar to the static model, in the first period the firm has three kinds of strategies: offering the high-quality product to high-type consumers only (T_1), offering the high-quality product to all consumers (T_2), and offering high quality to high-type consumers and low quality to low-type consumers (T_3). We notice that, by adopting strategies T_1 and T_3, the firm fully learns consumer types from their purchase history. Hence, in the second period consumers are separated into two segments, and the firm can and will offer the two segments the same product but at different prices. In this case, the firm is going to offer only the high-quality product $\overline{q}$ in the second period, but charge $u(\overline{q}, \theta_H)$ for those customers who purchased $\overline{q}$ in the first period and $u(\overline{q}, \theta_L)$ otherwise. Foreseeing this outcome, a high-type customer knows that he is going to get no consumer surplus in the second period if he reveals his type in the first period. Hence, high-type customers are willing to reveal their types only if the monopolist compensates them for the second-period information rent beforehand. On the other hand, strategy T_2 provides no information about consumer types; hence, in the second period the firm's maximization problem goes back to the static problem. To determine which of these three strategies is optimal, the magnitudes of α and the discount factor δ play important roles. First, we define two cutoff points of δ : [9]

$$
\begin{aligned}
\delta^* &= 1 - \frac{u(\underline{q}, \theta_H) - u(\underline{q}, \theta_L)}{u(\overline{q}, \theta_H) - u(\overline{q}, \theta_L)} \\
\delta^c &= \frac{\alpha S(\overline{q}, \theta_H) - S(\overline{q}, \theta_L)}{\alpha S(\overline{q}, \theta_H) - S(\overline{q}, \theta_L) - [\alpha S(\underline{q}, \theta_H) - S(\underline{q}, \theta_L)]},
\end{aligned}
$$

We then define three areas on the parameter space for (α, δ):

$$
\begin{aligned}
R_1 &= \{(\alpha, \delta) \in (0,1)^2 | \alpha \geq \Lambda(\underline{q}),\ or\ \Delta \leq \alpha \leq \Lambda(\underline{q})\ and\ \delta^* \leq \delta \leq \delta^c\} \\
R_2 &= \{(\alpha, \delta) \in (0,1)^2 | \alpha \leq \Delta,\ or\ \Delta \leq \alpha \leq \Lambda(\underline{q})\ and\ \delta \geq \delta^c\} \\
R_3 &= \{(\alpha, \delta) \in (0,1)^2 | \Delta \leq \alpha \leq \Lambda(\underline{q})\ and\ \delta \leq \delta^*\}
\end{aligned}
$$

Also let p_1^q denote the first-period price for the product q. The next proposition characterizes the monopolist's optimal first-period strategy with exogenously given qualities.[10]

Proposition 2 *In the dynamic model with two consumer types, the monopolist's optimal first-period strategy is characterized as follows:*

- *Case 1.* $\Lambda(\overline{q}) > \Lambda(\underline{q})$. *The monopolist will offer only the high-quality product* $\overline{q}$. *The price is given by* $p_1^{\overline{q}} = u(\overline{q}, \theta_H) - \delta[u(\overline{q}, \theta_H) - u(\overline{q}, \theta_L)]$ *and only high-type consumers purchase if* $\alpha \geq \Lambda(\overline{q})$, *and* $p_1^{\overline{q}} = u(\overline{q}, \theta_L)$ *and both types purchase otherwise.*

- *Case 2.* $\Lambda(\overline{q}) \leq \Lambda(\underline{q})$. *In this case,* $\Delta \leq \Lambda(\overline{q}) \leq \Lambda(\underline{q})$ *and the monopolist will offer (i) only* $\overline{q}$ *at* $p_1^{\overline{q}} = u(\overline{q}, \theta_H) - \delta[u(\overline{q}, \theta_H) - u(\overline{q}, \theta_L)]$ *and only high-type consumers purchase if* $(\alpha, \delta) \in R_1$, *(ii)only* $\overline{q}$ *at the price* $p_1^{\overline{q}} = u(\overline{q}, \theta_L)$ *and all consumers purchase if* $(\alpha, \delta) \in R_2$, *and (iii) both*

[9] δ^c is not necessarily between zero and one. See the appendix for more details.

[10] The optimal second-period strategy can be readily seen once we know the first-period strategy.

qualities with $p_1^{\underline{q}} = u(\underline{q}, \theta_L)$ *and* $p_1^{\overline{q}} = u(\overline{q}, \theta_H) - u(\underline{q}, \theta_H) + u(\underline{q}, \theta_L) - \delta[u(\overline{q}, \theta_H) - u(\overline{q}, \theta_L)]$ *if* $(\alpha, \delta) \in R_3$.

Proof. See the Appendix.

Several points deserve note here. First, if we ignore the condition for δ and focus on the range of α, we observe that the static and dynamic models have exactly the same cutoff points for α in determining the optimality among these three strategies! This coincidence, however, occurs for different reasons. We should be aware that in the dynamic model without commitment, the optimal pricing strategy is not just simply a duplicate of the static optimal pricing strategy. In the dynamic model, strategies T_1 and T_3 give us full information about consumer types in the second period, but the firm must pay an information rent to the high type in the first period. Hence, intrinsically the comparison between these two strategies in the dynamic model is the same as that in the static model. Strategy T_2, however, is different from the other two in the sense that it doesn't try to solicit consumers to reveal their types. Hence, T_2 need not pay the information rent in the first period, and, in contrast to T_1 and T_3, the second-period optimal strategy following T_2 is simply a duplicate of the optimal static pricing strategy. Therefore, the comparison between T_2 and the other two strategies in the dynamic model is different from the comparison in the static model. Even though both models share the same cutoff points of α in determining optimality, the contracts are different.

Second, by incorporating the effect of δ, we observe that in the dynamic model strategy T_1 or T_2 is more likely to be optimal than in the static model. In the static model, T_3 is the optimal strategy if α is in between ($\Delta \leq \alpha \leq \Lambda(\underline{q})$). If the seller-buyer relationship is dynamic and the firm is not able to commit to its behavior in the future, then T_1 or T_2 could be optimal when $\Delta \leq \alpha \leq \Lambda(\underline{q})$ provided that δ is not too small (i.e., consumers are not impatient).

When the increasing percentage differences condition holds and α is in between ($\Delta \leq \alpha \leq \Lambda(\underline{q})$), consumer patience δ determines the optimality among these three strategies. If the firm wants to separate consumers by inducing high-type consumers to take the high-quality product and low-type consumers to take the low-quality product, then in the first period it needs to pay high-type consumers not only the first-period information rent but also the second-period information rent. When consumers are more patient, the discounted information rent is higher, and separating types becomes more difficult. It can be shown that when $\delta > \delta^*$ the information rent is too high and the incentive constraint can never be met. Hence, strategy T_3 (offering two qualities in the first period) is achievable only if $\delta \leq \delta^*$. If this condition holds, then by reasoning similar to that discussed in the static model, T_3 is indeed optimal. Thus the firm will offer two qualities in the first period if and only if the percentage change in social surplus from product upgrades is increasing in the consumer type, the fraction of high types in the population is in between and consumers are impatient. On the other hand, if $\delta > \delta^*$, then T_3 is no longer achievable, and the firm has to decide to serve all consumers or only high-type consumers in the first period. If the firm serves only high-type consumers in the first period, then it has to pay the second-period information rent to high-type consumers ahead of time. However, if consumers are quite patient ($\delta \geq \delta^c$), then the discounted information rent is too high to be worthwhile for the firm, and such a strategy, although it is still achievable, is not profitable. Hence, the firm will choose to cut the price and serve the whole market in the first period. On the contrary, if consumer patience is in between ($\Delta \leq \alpha \leq \Lambda(\underline{q})$), then the firm will choose to charge a higher price and serve only high-type consumers in the first period and extract consumer surplus from all consumers in the second period.

Third, behavior-based price discrimination does not occur when the fraction of high types in the population is small or the increasing percentage differences condition holds but consumers are quite patient ($\delta \geq \delta^c$). There are two first-period strategies that allow the firm to price discriminate consumers based on their purchase history: offer the high quality product to high-type consumers only (T_1) or offer different qualities to different consumers (T_3). When the fraction of high types in the population is small, however, either strategy sacrifices too much profit from low-type consumers; hence, neither strategy is more profitable than T_2. When the increasing percentage differences condition holds but consumers are quite patient ($\delta \geq \delta^c$), T_3 is not feasible and T_1 is dominated by T_2 as explained before. Hence, the firm will forego the opportunity to discriminate consumers according to their purchase history in these two situations.

Finally, the firm will offer two versions of the product to the market in either period if and only if the increasing percentage differences condition holds and consumers are impatient or quite patient. Moreover, the firm will never offer two versions of the product in *both* periods. If the increasing percentage differences condition holds, then the firm would like to offer two qualities in the first period, but this strategy is not feasible when consumers are quite patient. Hence, in this case the firm will offer the high-quality product to all consumers in the first period and postpone offering both qualities to the second period. If the increasing percentage differences condition holds and consumers are impatient, then the firm will offer both qualities in the first period and offer only the high-quality product in the second period.

Corollary 1 *(i) If $\Lambda(\overline{q}) > \Lambda(\underline{q})$, then the firm will offer only one quality in the first period. (ii) Suppose $\Lambda(\overline{q}) \leq \Lambda(\underline{q})$, then the first table below describes the relationship between optimal quality provision in the first-period and the values of α and δ, and the execution of behavior-based price discrimination is summarized in the second table.*

First period optimal quality provision			
	$\alpha \leq \Delta$	$\Delta \leq \alpha \leq \Lambda(\underline{q})$	$\alpha \geq \Lambda(\underline{q})$
$\delta \leq \delta^*$	$\overline{q}$ to all	both qualities	$\overline{q}$ to high
$\delta^* \leq \delta \leq \delta^c$	$\overline{q}$ to all	$\overline{q}$ to high	$\overline{q}$ to high
$\delta \geq \delta^c$	$\overline{q}$ to all	$\overline{q}$ to all	$\overline{q}$ to high

Behavior-Based Price Discrimination			
	$\alpha \leq \Delta$	$\Delta \leq \alpha \leq \Lambda(\underline{q})$	$\alpha \geq \Lambda(\underline{q})$
$\delta \leq \delta^*$	N	Y	Y
$\delta^* \leq \delta \leq \delta^c$	N	Y	Y
$\delta \geq \delta^c$	N	N	Y

What if the firm is able to commit to its second-period behavior? In this case, it is easy to see that we have the following result:

Proposition 3 *In the dynamic model with commitment, the monopolist's optimal pricing strategy is to repeat the optimal static pricing strategy in each period.*

Hence, the firm finds it optimal to commit not to use the purchase history information to price discriminate consumers, a common result in the literature on dynamic contracting without commitment or BBPD.[11] Acquisti and Varian (2005) study behavior-based price discrimination with only one quality and no production cost. Since there is only one quality in their model, the firm can only do BBPD by observing if a consumer made a purchase in the first period or not. They ask whether a monopolist with the ability to commit can do better than offering a flat price in each period through some form of conditioning pricing and reach the same conclusion as we obtain here.[12] Here we consider a more general model in which the firm is able to offer multiple qualities; hence, it is not surprising to see that the firm can do better in our model as it can still extract some consumer surplus from the low types by offering a low quality product.

Since Proposition 3 applies to the situation when the firm is not able to track consumers' purchase history as well, it can be readily seen from Propositions 2 and 3 that in most cases being able to track purchase history benefits consumers when the firm cannot commit, as the firm with knowledge of consumer types is more likely to offer the more efficient product $\overline{q}$ and cover more consumers in the second period. The only exception is the case when $\Delta \leq \alpha \leq \Lambda(\underline{q})$ and $\delta^* \leq \delta \leq \delta^c$. In this case compared to the case when the firm cannot track purchase history, the high-type consumers are hurt because they can only get the second-period information rent.

2.3 Optimal Product Line Design Dynamics

Next we extend the basic model by allowing the monopolist to choose the product quality optimally. Suppose now the firm can choose any quality $q \in [0, 1]$ with the unit production cost $c(q)$. Again, let p_1^q denote the price of the product with quality q offered by the monopolist in the first period. First we adopt the following assumption:

Assumption 2 *(i) $u_q(q, \theta) > 0$, $u(q, \theta)$ and $u_q(q, \theta)$ are increasing in θ, (ii) $S_q(1, \theta) \geq 0$, (iii) $S(0, \theta) \leq 0$ and $S(q, \theta) > 0$ for all θ and some $q \in (0, 1)$, and (iv) $S_{qq}(q, \theta) < 0$ and $u_{qq}(q, \theta)$ is weakly increasing in θ.*

(i) is the single-crossing condition. (ii) assumes that the quality constraint $q \leq 1$ is weakly binding. (iii) guarantees that the firm will choose positive qualities, and (iv) is a sufficient condition for the q^* defined below to be unique. Define

$$q^* = \arg\max_{q \in [0,1]} \Gamma(q) = S(q, \theta_L) - \alpha S(q, \theta_H)$$

$$I(q) = \left(\frac{S_q(1, \theta_L)}{S_q(1, \theta_L)}, \frac{S(q, \theta_L)}{S(q, \theta_H)}\right)$$

Anderson and Dana (2006) establish the following result for the static model:

Proposition 4 *(Anderson and Dana (2006))In the static model, the monopolist's optimal strategy is to offer (i) a single quality $q = 1$ at the price $p = u(1, \theta_H)$ if $\alpha \geq \max\{\Lambda(1), \Lambda(q^*)\}$, (ii) a single quality $q = 1$ at the price $p = u(1, \theta_L)$ if $\alpha \leq \min\{\Lambda(1), \frac{S_q(1,\ \theta_L)}{S_q(1,\ \theta_H)}\}$, and (iii) two qualities $q = 1$ and $q = q^*$ with $p(q^*) = u(q^*, \theta_L)$ and $p(1) = u(q^*, \theta_L) + u(1, \theta_H) - u(q^*, \theta_H)$ if $\alpha \in I(q^*)$.*

[11] See, for example, Acquisti and Varian (2005), Laffont and Tirole (1988) and Taylor (2004).

[12] If the firm cannot commit to its future actions, they show that in equilibrium consumers must play mixed strategies which do not appear in our model. The reason is that in their model $\delta = 1$, and here we focus our analysis on the case where $0 < \delta < 1$.

The intuition behind this result is similar to the case when qualities are exogenously given. When the fraction of high types in the population is either high or low, the firm will just offer one quality to the market. To the contrary, two qualities will be offered to the market if and only if two conditions hold: the interval $I(q^*)$ is not empty and $\alpha \in I(q^*)$. However, we notice that if $S(q,\theta)$ is log submodular, then $I(q^*)$ is empty. $I(q^*)$ is never empty if $S(q,\theta)$ is log supermodular. Hence, we have the following corollary:

Corollary 2 *In the static model, the firm will offer (i) only one quality if $S(q,\theta)$ is log submodular, and (ii) multiple qualities if $S(q,\theta)$ is log supermodular and $\alpha \in I(q^*)$.*

Recall that in the dynamic model with exogenously given qualities offering two qualities in the first period is not achievable when consumers are patient, i.e., $\delta > \delta^*$. This infeasibility is due to the fact that the information rent is too high to make the high quality product appealing to the high type. Hence, the firm is not able to price discriminate consumers by qualities in the first period. Now we allow the firm to choose the qualities optimally, which gives the firm more flexibility to price discriminate consumers as shown in the next proposition. To characterize the optimal first-period strategy in the dynamic environment, first we define

$$\begin{aligned}
\lambda(q) &= 1 - \frac{u(q,\theta_H) - u(q,\theta_L)}{u(1,\theta_H) - u(1,\theta_L)} \\
\xi(q) &= \alpha[S(1,\theta_H) - S(q,\theta_H)] - [S(1,\theta_L) - S(q,\theta_L)] \\
\tilde{\delta} &= \frac{\alpha S(1,\theta_H) - S(1,\theta_L)}{\xi(q^*)},
\end{aligned}$$

Since $\lambda(q)$ is decreasing in q, it has an inverse function denoted by $\lambda^{-1}(\cdot)$. To make the next proposition clear, we divide the parameter space for $(\alpha,\, \delta)$ into four regions: $\Theta = (0,1)^2 = \Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4$, where

$$\begin{aligned}
\Omega_1 &= \{(\alpha,\delta) \in (0,1)^2 | \\
\alpha &\geq \max[\Lambda(1), \Lambda(q^*)],\ \mathit{or}\ \lambda(q^*) \leq \delta \leq \min[\lambda(0), \tilde{\delta}]\ \mathit{and}\ \alpha \geq \Lambda(\lambda^{-1}(\delta)),\ \mathit{or}\ \lambda(0) \leq \delta \leq \tilde{\delta}\} \\
\Omega_2 &= \{(\alpha,\delta) \in (0,1)^2 | \\
\alpha &\leq \min[\Lambda(1), \frac{S_q(1,\theta_L)}{S_q(1,\theta_H)}],\ \mathit{or}\ \delta \geq \max[\tilde{\delta}, \lambda(q^*), \frac{\xi(\lambda^{-1}(\delta))}{\xi(q^*)}],\ \mathit{or}\ \delta \geq \max[\tilde{\delta}, \lambda(0)]\} \\
\Omega_3 &= \{(\alpha,\delta) \in (0,1)^2 | \alpha \in I(q^*)\ \mathit{and}\ \delta \leq \lambda(q^*)\} \\
\Omega_4 &= \{(\alpha,\delta) \in (0,1)^2 | \lambda(q^*) \leq \delta \leq \min[\lambda(0), \frac{\xi(\lambda^{-1}(\delta))}{\xi(q^*)}]\ \mathit{and}\ \alpha \leq \Lambda(\lambda^{-1}(\delta))\}.
\end{aligned}$$

Proposition 5 *In the dynamic model with endogenous qualities, the monopolist's optimal first-period strategy is to offer (i) a single quality $q = 1$ at the price $p = u(1,\theta_H) - \delta[u(1,\theta_H) - u(1,\theta_L)]$ and only high-type consumers purchase if $(\alpha,\delta) \in \Omega_1$, (ii) a single quality $q = 1$ at the price $p = u(1,\theta_L)$ and all consumers purchase if $(\alpha,\delta) \in \Omega_2$, (iii) both qualities $q = 1$ and $q = q^*$ with $p_1^{q^*} = u(q^*,\theta_L)$ and $p_1^1 = u(1,\theta_H) - u(q^*,\theta_H) + u(q^*,\theta_L) - \delta[u(1,\theta_H) - u(1,\theta_L)]$ if $(\alpha,\delta) \in \Omega_3$, (iv) both qualities $q = 1$ and $q = \lambda^{-1}(\delta)$ with $p_1^{\lambda^{-1}(\delta)} = u(\lambda^{-1}(\delta),\theta_L)$ and $p_1^1 = u(1,\theta_H) - u(\lambda^{-1}(\delta),\theta_H) + u(\lambda^{-1}(\delta),\theta_L) - \delta[u(1,\theta_H) - u(1,\theta_L)]$ if $(\alpha,\delta) \in \Omega_4$.*

Proof. See the Appendix.

Basically this result is the same as what we obtain in the model with exogenously given qualities: only one quality will be offered in the first period when consumers are patient or α is either high or low. The condition for δ is slightly more complicated. If the restriction on δ can be ignored, then the firm's optimal strategy is to offer two qualities $q = 1$ and $q = q^*$ in the first period to price discriminate consumers when $\alpha \in I(q^*)$. This strategy is not achievable when consumers are patient ($\delta > \lambda(q^*)$), since the high type can get more surplus by pretending he is of low type. One way to avoid this problem is to make the low quality product unattractive to the high type. To do this, the firm can degrade the quality of the second product ($q = 1$ for the first product) to $q = \lambda^{-1}(\delta) < q^*$. This strategy could be optimal provided that the profit sacrificed from the low type is small. In this case, the quality gap $(1 - \lambda^{-1}(\delta))$ is increasing in consumer patience.[13] On the other hand, if consumers are patient enough, then offering two qualities in the first period is either unprofitable or not achievable. Also it is not hard to verify that $\Omega_3 \cup \Omega_4$ is empty if $S(q, \theta)$ is log submodular. The following corollary summarizes these results.

Corollary 3 *In the dynamic model with endogenous qualities, the monopolist will offer only one quality in the first period if (i) the social surplus function $S(q, \theta)$ is log submodular, or (ii) $\delta > \lambda(0)$, or (iii) $\delta \leq \lambda(0)$ but $\delta > \max\{\lambda(q^*), \frac{\xi(\lambda^{-1}(\delta))}{\xi(q^*)}\}$.*

3 The Continuous Consumer Type Model

3.1 The Static Model

This subsection introduces the static model with endogenous qualities when the space of consumer types is a continuum and summarizes the main results obtained by Anderson and Dana (2006). The dynamic version of the model will be studied in the next subsection. There are two main differences between the discrete model and the continuous model. First, in contrast to the discrete model in which full separation with different types choosing different qualities could be optimal for some parameters (Proposition 4 (iii) and (iv)), we show that full separation can never be achievable when the consumer type space is a continuum (Proposition 6). Hence, some pooling is inevitable for almost all types when the consumer types are continuous. Second, as can be readily seen in Proposition 4, in an equilibrium of the discrete type model consumers may *upgrade* the product quality in the second period, but they never choose to *downgrade*. In the continuous model, on the other hand, we show that there are some equilibria in which a positive measure of consumers choose to downgrade the product quality in the second period (Proposition 8). The intuition for these results will be explored further in the next subsection.

Assume that the consumers' types are distributed on $[\underline{\theta}, \overline{\theta}]$ with strictly positive density $f(\theta)$ and cumulative distribution $F(\theta)$. We define a *contract* as a price-quality pair (p, q) that specifies the quality q a consumer can get at price p. Assume $q \in [0, 1]$. With a slight abuse of the notation, we use $p(q)$ to represent the price in the pair (p, q). Let $\mathcal{C} \subset \mathbb{R}_+ \times [0, 1]$ denote the menu of

[13] As pointed out by Jan Bouckaert, δ can be thought of as the probability of repurchase, in which case our result implies that if it is optimal for the firm to offer two qualities in the first period, then the quality gap is higher when the probabililty of repurchase is higher.

contracts offered by the monopolist. Consumer θ chooses a contract in $\mathcal{C}$ to maximize the consumer surplus $u(q,\theta)$ less the price p. $u(q,\theta)$ satisfies $u_\theta(q,\theta) > 0$, $u_q(q,\theta) > 0$, and $u_{q\theta}(q,\theta) > 0$ on $[0,1] \times [\underline{\theta}, \overline{\theta}]$. For a given $\mathcal{C}$, define $(p(q(\theta)), q(\theta))$ to be a contract that maximizes consumer $\theta's$ utility, i.e., $(p(q(\theta)), q(\theta)) \in \arg\max_{(p,q)\in\mathcal{C}} u(q,\theta) - p$. Consumer θ makes a purchase if and only if the individual rationality constraint $u(q(\theta),\theta) - p(q(\theta)) \geq 0$ is met. The firm's unit cost of producing quality q is $c(q)$. Again, let $S(q,\theta) = u(q,\theta) - c(q)$ denote the surplus function. Define $\phi(\theta) = \arg\max_{q\in[0,1]} S(q,\theta)$ and ζ to be the point such that $S(\phi(\zeta),\zeta) = 0$. The firm's problem is to choose a menu $\mathcal{C} \subset \mathbb{R}_+ \times [0,1]$ to maximize its profit $\int_{\theta_l}^{\overline{\theta}} [p(q(\theta)) - c(q(\theta))] dF(\theta)$, where the cutoff point θ_l solves $u(q(\theta_l),\theta_l) - p(q(\theta_l)) = 0$. By invoking the direct revelation principle and integration by parts, the firm's problem is equivalent to picking θ_l and a nondecreasing quality function $q(\cdot) : [\theta_l, \overline{\theta}] \to [0,1]$ to

$$\max \int_{\theta_l}^{\overline{\theta}} [S(q(\theta),\theta) - \frac{1-F(\theta)}{f(\theta)} S_\theta(q(\theta),\theta)] dF(\theta).$$

We make the following assumption:[14]

Assumption 3 *(i) The hazard rate* $\frac{1-F(\theta)}{f(\theta)}$ *is decreasing, (ii)* $\zeta \in (0,1)$, *(iii)* $S(0,\theta) < 0 \ \forall\theta \in [\underline{\theta}, \overline{\theta}]$, *and (iv)* $S_{qq}(q,\theta) < 0$, $S_{qq\theta}(q,\theta) \geq 0$, *and* $S_{q\theta\theta}(q,\theta) \leq 0$.

Condition (ii) implies that the firm will never choose to cover the whole market. (iii) implies that the firm will never choose $q = 0$, and (i) and (iv) allow us to ignore the monotonicity constraint in q and solve the problem pointwisely. Anderson and Dana establish the following result:

Proposition 6 *(Anderson and Dana (2006))(i) If* $S(q,\theta)$ *is log submodular* $\forall(q,\theta) \in (0,1) \times [\underline{\theta}, \overline{\theta}]$, *then the firm's optimal strategy is to offer a single contract with* $q = 1$. *(ii) If* $S(q,\theta)$ *is log supermodular at* $(1,\theta)$ $\forall\theta \in [\underline{\theta}, \overline{\theta}]$, *then the firm's optimal strategy is to offer multiple contracts.*

This proposition provides an elegant condition to determine whether price discrimination through offering multiple qualities is profitable. The condition states that if the percentage change in social surplus from product upgrades is decreasing in the consumer type, then the firm finds it optimal to offer only one contract. On the other hand, if the percentage change in social surplus from product upgrades is increasing in the consumer type, then the firm can increase its profit by offering multiple qualities. Contrary to the discrete model in which the fraction of high types plays an important role in determining the optimal strategy, log supermodularity alone is a sufficient condition for offering multiple contracts to be optimal.

3.2 The Properties of Dynamic Optimal Contracts

We now start our analysis of the two-period version of the model with customer recognition. The timing of the game is as follows: (i) In the first period, the firm chooses a menu of contracts $\mathcal{C}_1 \subset \mathbb{R}_+ \times [0,1]$, and then each consumer makes his choice. (ii) Consumers are divided into several

[14](iv) is a sufficient second-order condition guaranteeing that our maximization problem is well-defined and the optimal quality function $q(\cdot)$ is strictly increasing or equal to one. It can be replaced by the following slightly weaker condition: $S_{qq} - \frac{1-F}{f} S_{qq\theta} < 0$ and $S_{q\theta\theta} < [\frac{2f}{1-F} + \frac{f'}{f}] S_{q\theta}$.

purchasing segments based on the contract they took in the first period. The firm provides a menu of contracts to each group. (iii) Finally, each consumer makes his second-period choice. Let $(p_1(\theta), q_1(\theta)) \in \mathcal{C}_1$ denote the contract taken by consumer θ in the first period. Define $A(\theta) = \{\theta' \in [\underline{\theta}, \overline{\theta}] | (p_1(\theta'), q_1(\theta')) = (p_1(\theta), q_1(\theta))\}$, the set of types that choose the same contract as type θ in the first period, and $N = \{\theta \in [\underline{\theta}, \overline{\theta}] |\ \theta$ does not take any contract in the first period$\}$. Then $\{N\} \cup \{A(\theta)\}_{\theta \in [\underline{\theta}, \overline{\theta}]}$ is a partition of the type space $[\underline{\theta}, \overline{\theta}]$. Let $\mathcal{C}_{2A} \subset \mathbb{R}_+ \times [0, 1]$ denote the set of contracts offered by the firm to group A in the second period and $p_2(\theta|A), q_2(\theta|A)) \in \mathcal{C}_{2A}$ the contract taken by consumer $\theta \in A$ in the second period.[15] The monopolist acts to maximize the discounted value of profits, using the common discount factor $0 < \delta < 1$. Consumer $\theta's$ behavior is to maximize the discounted sum of per-period utilities. To solve this problem, we use perfect Bayesian equilibrium (PBE) as the solution concept. For a given $\mathcal{C}_1$, a continuation equilibrium is a PBE for the subgame following $\mathcal{C}_1$. Next, we give a definition and prove the nonexistence of fully separating continuation equilibria.

Definition 1 *For a given $\mathcal{C}_1$, a continuation equilibrium with the partition $\{N\} \cup \{A(\theta)\}_{\theta \in [\underline{\theta}, \overline{\theta}]}$ is said to be fully separating if $N = [\underline{\theta}, a]$ for some $a \in [\underline{\theta}, \overline{\theta})$ and $A(\theta)$ is a singleton for all $\theta \in (a, \overline{\theta}]$.*[16]

Proposition 7 *For any given $\mathcal{C}_1$, there exists no fully separating continuation equilibrium.*

Proof. Suppose to the contrary that there exists a fully separating continuation equilibrium with $N = [\underline{\theta}, a]$ for some $a \in [\underline{\theta}, \overline{\theta})$. Let $(p_1(q_1(\theta)), q_1(\theta)) \in \mathcal{C}_1$ denote the contract taken by consumer θ in the first period, where $\theta \in (a, \overline{\theta}]$. WLOG we assume $a \geq \zeta$. Since each consumer $\theta \in (a, \overline{\theta}]$ fully reveals his type in the first period, the firm will extract all surplus from them in the second period. Hence in the second period the firm will offer $(u(\phi(\theta), \theta), \phi(\theta))$ to the consumer who took $(p_1(q_1(\theta)), q_1(\theta))$ in the first period, and the value function is $V(\theta) = u(q_1(\theta), \theta) - p_1(q_1(\theta)) + \delta[u(\phi(\theta), \theta) - u(\phi(\theta), \theta)] = u(q_1(\theta), \theta) - p_1(q_1(\theta))$. The incentive constraint implies that $p_1(q_1(\theta))$ and $q_1(\theta)$ are monotone increasing and hence, differentiable a.e. Pick any θ^* at which $V(\cdot)$ is differentiable. Then $V'(\theta^*) = u_q(q_1(\theta^*), \theta^*)q_1'(\theta^*) + u_\theta(q_1(\theta^*), \theta^*) - p_1'(q_1(\theta^*))q_1'(\theta^*) = u_\theta(q_1(\theta^*), \theta^*)$, where $u_q(q_1(\theta^*), \theta^*) - p_1'(q_1(\theta^*)) = 0$ by the incentive constraint again. On the other hand, applying the general envelope theorem (Theorem 2, Milgrom and Segal (2002)) gives us $V'(\theta^*) = u_\theta(q_1(\theta^*), \theta^*) + \delta u_\theta(\phi(\theta^*), \theta^*)$, a contradiction. *Q.E.D.*

Contrary to the static model with continuous consumer types in which full separation is achievable and may be profitable, Proposition 7 shows that full separation is never achievable in the dynamic contracting problem without commitment. This outcome can be explained by the well-known "*ratchet* (*bunching*) *effect*."[17] A high-type consumer will face an unfavorable contract in the second period if he reveals his type in the first period by taking a high quality product Therefore, he is disinclined to reveal his type early in the dynamic relationship.

[15] Without loss of generality here we assume that the exit option is in all menus $\mathcal{C}_1$ and $\mathcal{C}_{2A}$. Consumers can always choose to not take any contract.

[16] To avoid the issue of determining the cutoff point, Laffont and Tirole (1988) assume that the social utility of the project is large enough so that it is always worth carrying out the project, meaning that $N = \varnothing$ in their model. Here we always have an interior cutoff point, hence the "pooling part" N is inevitable in defining fully separating.

[17] See Laffont and Tirole (1988) and Weitzman (1980).

To the best of our knowledge, no one has been able to characterize the whole set of incentive contracts and solve for the optimal dynamic contracts when the type space is a continuum and the principal cannot commit to his future actions, for there are too many possible partitions on the type space given any first-period menu of contracts and the revelation principle fails in dynamic contracting without commitment.[18] Laffont and Tirole (1988) encounter the same problem when they study cost regulation in the dynamic framework, and they restrict their analysis to the properties of continuation equilibria. Moreover, they give an example to show that there exists a continuation equilibrium which exhibits the infinite reswitching property (hence it is nonpartitional). As suggested by Laffont and Tirole, however, it is natural for us to study the so-called *partitional* continuation equilibrium.

Definition 2 *A continuation equilibrium for a given* $\mathcal{C}_1$ *is said to be partitional if the number of segments on the type space* $\{N\} \cup \{A(\theta)\}_{\theta \in [\underline{\theta}, \overline{\theta}]}$ *induced by* $\mathcal{C}_1$ *is countable and each segment is connected.*

A partitional continuation equilibrium rules out the situation in which there is a contract in $\mathcal{C}_1$ taken by some low-type and high-type consumers, but not taken by those in between. It is actually a dynamic version of the single crossing condition in the sense that each pair of the two-period utility functions induced by $\mathcal{C}_1$ and the firm's beliefs can cross at most once. The set of all PBE with a partitional continuation equilibrium is said to be the class of partitional PBE. According to this definition, if a PBE with the first-period menu $\mathcal{C}_1$ has a partitional continuation equilibrium, then there can be at most countably many contracts in $\mathcal{C}_1$ taken by consumers. However, the following proposition tells us that the number of feasible first-period contracts can be substantially trimmed in equilibrium.

Proposition 8 *In a PBE with partitional continuation equilibrium, if* (p_1, q_1) *and* (p_2, q_2) *in* $\mathcal{C}_1$ *are taken by some consumers, then* $|q_1 - q_2| \geq \delta \min_{\theta \in [\zeta, \overline{\theta}]} \frac{u_\theta(\phi(\theta), \theta)}{u_{\theta q}(1, \theta)}$.

Proof. See the Appendix.

Within the class of partitional PBE, Proposition 8 demonstrates the extensive influence of the *ratchet effect*. The ratchet effect expresses the unwillingness of the high types to reveal their types by taking a high quality product. A high-type consumer is inclined to take a high quality product only if the quality difference is sufficiently large so that the utility gain from taking a high quality product is larger than the discounted information rent he can acquire by taking a low quality product in the first period instead. Accordingly Proposition 8 provides a lower bound for the quality difference taken by consumers in the first period in equilibrium. It states that the difference between the two first-period qualities in a PBE with partitional continuation equilibrium cannot be smaller than the discount factor times some constant that is determined by the structure of the social surplus function.[19] A higher discount factor implies a higher discounted information rent;

[18] With the assumption that there are only finitely many types of agents, Bester and Strausz (2001) study the optimal contracting problem for environments in which the principle cannot fully commit to the outcome induced by the mechanism and derive a revised version of the revelation principle in the sense that each agent reports his true type with some positive probability which is not necessary equal to one. They then reduce the original question to a standard programming problem by applying Caratheodory's theorem.

[19] Recall that ζ is determined by $S(\phi(\zeta), \zeta) = 0$.

therefore the lower bound is greater. $\frac{u_{\theta q}}{u_\theta}$ measures the percentage change in the marginal utility with respect to θ from product upgrades. A higher $\frac{u_{\theta q}}{u_\theta}$ implies a higher utility gain from taking a high quality product and consequently a smaller lower bound. In addition, since the quality offering is bounded between 0 and 1, this proposition is equivalent to saying that in a PBE with partitional continuation equilibrium the number of relevant first-period contracts cannot be greater than the inverse of this lower bound.[20] Hence, we establish an upper bound for the number of product versions offered in the first period, which greatly simplifies the procedure for searching for the optimal dynamic contracts.

Corollary 4 *In a PBE with partitional continuation equilibrium, the number of contracts taken by consumers in the first period cannot be greater than* $1/[\delta \min_{\theta\in[\zeta,\overline{\theta}]} \frac{u_\theta(\phi(\theta),\theta)}{u_{\theta q}(1,\theta)}]$.

Proposition 7 demonstrates that some pooling must occur in equilibrium for almost all types, but it does not tell us how much pooling is required. This corollary supplements Proposition 7 by explicitly telling us the maximum number of first-period contracts that can be allowed in a PBE with partitional continuation equilibrium. With this corollary in hand and confining ourselves to the class of partitional PBE, characterizing the optimal dynamic contracts becomes an extremely complicated but theoretically solvable maximization problem. To see this, we first observe that if the firm offers a menu of n contracts in the first period which induces a partitional continuation equilibrium, then the segments $\{N\} \cup \{A(\theta)\}_{\theta\in[\underline{\theta},\overline{\theta}]}$ can be fully characterized by an increasing sequence of cutoff points $\{\theta_i\}_{i=1}^n$ on $[\underline{\theta},\overline{\theta}]$. The optimal strategy in each segment can be specified using standard techniques, and the optimal dynamic strategy with n first-period contracts can be solved by choosing $\{\theta_i\}_{i=1}^n$ appropriately. Doing this for menus of n contracts for all n below the upper bound, the optimal first-period menu of contracts is the one that generates the most profit. However, we are interested here in when the firm finds it optimal to offer only one quality in the first period and what this equilibrium looks like. The answers are provided in the next two propositions.

Remember that in the static model the firm's problem is to choose a cutoff point θ_l and a nondecreasing quality function $q(\cdot) : [\theta_l, \overline{\theta}] \to [0,1]$ to maximize $\int_{\theta_l}^{\overline{\theta}}[S(q(\theta),\theta) - \frac{1-F(\theta)}{f(\theta)}S_\theta(q(\theta),\theta)]dF(\theta)$. Let θ_l^* denote the optimal cutoff point in the static model. To simplify the expression, we define

$$\chi(\theta) = S(\phi(\theta),\theta) - \frac{1-F(\theta)}{f(\theta)}S_\theta(\phi(\theta),\theta),$$

the virtual profit evaluated at type θ and quality $\phi(\theta)$. Our next proposition provides some interesting properties of the optimal contracts over two periods when offering a single contract in the first period is optimal for the firm. Let θ_{1l}^* denote the optimal first-period cutoff point, θ_N^* the optimal cutoff point in N, $A = [\theta_{1l}^*, \overline{\theta}]$ the set of consumers who purchased in the first period, and q_1^* the optimal first-period quality.

Proposition 9 *In the dynamic model, suppose that the firm's optimal strategy is to offer a single contract in the first period and that this contract induces a partitional continuation equilibrium. Then we have (i) (optimal cutoff points)* $\theta_{1l}^* > \theta_l^* > \theta_N^*$, *(ii) (quality provision dynamics)* $q_2(\theta_{1l}^*|A) \leq q_1^* = q_2(\theta_{1l}^*|N) = \phi(\theta_{1l}^*)$, *and (iii) (comparative statics) sign* $\frac{\partial\theta_{1l}^*}{\partial\delta}$ = *sign* $\chi(\theta_{1l}^*)$ *and sign* $\frac{\partial q_1^*}{\partial\delta}$ = *sign* $\chi(\theta_{1l}^*)$ $\forall q_1^* \neq 1$.

[20] Assuming the utility function is quadratic, Laffont and Tirole (1988) provide a lower bound on the difference between the first-period cost levels. Our result holds for any utility function satisfying assumption 3.

Proof. See the Appendix.

Three remarks are noteworthy. First, we observe that when facing a dynamic relationship and offering a single contract in the first period is optimal, the firm chooses to serve fewer consumers than it does in the static setup ($\theta_{1l}^* > \theta_l^*$) in the first period. This result is not as straightforward as it seems. A rough intuition behind this result is the following. If the firm does business with consumers only once, then it is going to serve all consumers down to the type that gives it zero virtual profit, i.e., θ_l^* is determined by equating the social surplus to information rent. On the other hand, in the repeated seller-buyer relationship with customer recognition, the change of θ_{1l} affects not only the first-period profit but also the second-period profit in segments A and N. To be precise, the first-period cutoff point θ_{1l} plays the role of balancing the first-period information rent effect against the ratchet effect.[21] Starting at $\theta_{1l} = \theta_l^*$, let us look at the effect of a change in θ_{1l}. Increasing the cutoff point from θ_l^* slightly has no effect on the profit in A since the maximization problem in A is the same as the static one and the virtual profit is zero at θ_l^*. The effects on the first-period profit and the profit in N are intertwined and somewhat more subtle. By increasing the cutoff point from θ_l^* slightly, the firm chooses to postpone serving consumer θ_l^* to the second period, sacrificing the virtual profit at θ_l^* in the first period but recouping it in the second period. Recognizing that the virtual profit at θ_l^* is zero when the firm chooses multiple qualities optimally, we conclude that the virtual profit at θ_l^* is negative when only one quality is offered as in our case, and, therefore, the net effect is positive. Combining this result and the fact that increasing θ_{1l} from θ_l^* slightly has no effect on the profit in A, the firm finds it optimal to serve fewer consumers in the first period than it does in the static relationship.

If the firm is able to commit to its future actions, then its optimal strategy is to offer twice the optimal static menu of contracts. Thus this result tells us that without commitment the firm chooses to cover fewer consumers in the first period. However, we notice that more consumers are covered in the second period as $\theta_l^* > \theta_N^*$. Therefore, it is not clear whether commitment increases social welfare or not.

The second observation concerns the quality provision dynamics. After θ_{1l} has been chosen, q_1 is simply set to maximize the social surplus at θ_{1l}, i.e., $q_1^* = \phi(\theta_{1l}^*)$. The standard "no distortion at the top" property gives us $q_2(\theta_{1l}^*|N) = \phi(\theta_{1l}^*)$ as well. The information rent effect distorts the quality downward; thus $q_2(\theta_{1l}^*|A) \leq \phi(\theta_{1l}^*)$. These results tell us that in the second period the firm will offer qualities no greater than the first-period quality q_1^* to those consumers who didnot make a purchase in the first period (low types), but may offer some qualities lower than q_1^* to those who made a purchase in the first period (high types). As such, there are some consumers choosing to "*downgrade*" the product in the second period, a phenomenon that has never been addressed in the literature.[22] Obviously the ratchet effect is the reason for this outcome. Because of the ratchet effect, pooling is unavoidable and accordingly the quality is distorted downward in the second period. While there is no quality distortion at θ_{1l}^* in the first period, we must have some people

[21] The information rent is defined as the gains accruing to a consumer solely from possessing payoff-relevent information. Hence, the information rent persists even when only one quality is offered in the first period.

[22] This paper focuses its analysis on nondurable goods. Fudenberg and Tirole (1998) study upgrades and buy-backs for a durable good. Since they assume that only a high-quality generation arrives in the second period, there is no possibility for consumers to downgrade the product. One possible and interesting extension of their basic model is to allow the firm to offer not only a high-quality version, but also an inferior or damaged version in the second period.

above θ_{1l}^* downgrading in the second period.[23]

Last but not least, we discuss the comparative statics of θ_{1l}^* and q_1^*. Roughly speaking, $\chi(\theta_{1l}^*)$ measures the first-period loss when the firm postpones serving consumer θ_l^* to the second period.[24] Positive $\chi(\theta_{1l}^*)$ implies a positive second-period gain from postponing serving consumer θ_l^* as the first-order effect vanishes at the optimal cutoff point θ_{1l}^*. Along with the increase of δ, the discounted second-period gain dominates the first-period loss, resulting the favor of increasing θ_{1l}^*. By the same token, a positive second-period loss is implied by a negative $\chi(\theta_{1l}^*)$. So if $\chi(\theta_{1l}^*)$ is negative, it is optimal for the firm to serve more customers in the first period as δ increases. On the whole, *sign* $\frac{\partial \theta_{1l}^*}{\partial \delta}$ = *sign* $\chi(\theta_{1l}^*)$. On the other hand, the first-period quality q_1^*, which is set to maximize the social surplus at θ_{1l}^*, is increasing in θ_{1l}^* by the single crossing property. Thus *sign* $\frac{\partial q_1^*}{\partial \delta}$ = *sign* $\chi(\theta_{1l}^*)$ as well.

Intuitively a monotone relationship between θ_{1l}^* and δ should be expected. The firm should be willing to cover more consumers in the first period when the future is less important to it. Since $\chi(\theta_{1l}^*)$ can be either positive or negative in equilibrium, however, θ_{1l}^* is not necessarily monotonically increasing in δ. Although the profit in N increases as θ_{1l} increases, the profit in A decreases at the same time. Thus the effect of θ_{1l} on the second-period profit is vague, and, as a result of this ambiguity, the effect of δ is indefinite. θ_{1l}^* is indeed monotonically increasing in δ under some situations. First, we notice that $\chi(\overline{\theta}) > 0$. This implies that $\eta = \sup\{\theta \in (\theta_l^*, \overline{\theta}) | \chi(\theta) \leq 0\} < \overline{\theta}$. So $\chi(\theta)$ is positive for all $\theta > \eta$, and consequently $\frac{\partial \theta_{1l}^*}{\partial \delta} > 0$ if $\theta_{1l}^* > \eta$. On the other hand, it is not difficult to see that $\chi(\theta) > 0$ for all $\theta \in (\theta_l^*, \overline{\theta})$ when $S(q, \theta)$ is log submodular. Therefore we have the following corollary:

Corollary 5 *Suppose that the firm's optimal strategy is to offer a single contract in the first period and that this contract induces a partitional continuation equilibrium. Let θ_{1l}^* denote the optimal first-period cutoff point. Then $\frac{\partial \theta_{1l}^*}{\partial \delta} > 0$ if $\theta_{1l}^* > \eta$ or if $S(q, \theta)$ is log submodular.*

It thus seems natural to ask the following question: When will the firm offer a single quality in the first period? The following proposition provides an answer to this question.

Proposition 10 *Within the class of partitional PBE, the firm's optimal strategy is to offer a single contract in the first period if (i) $S(q, \theta)$ is log submodular or (ii) $\delta > \frac{1}{\min_{\theta \in [\zeta, \overline{\theta}]} \frac{u_\theta(\phi(\theta), \theta)}{u_{\theta q}(1, \theta)}}$.*

Proof. See the Appendix.

Similar to the result in the model with two consumer types, the firm finds it optimal *not* to price discriminate consumers through qualities in the first period when the percentage change in social surplus from product upgrades is decreasing in the consumer type (i.e., $S(q, \theta)$ is log submodular). On the other hand, when $S(q, \theta)$ is not log submodular, offering multiple qualities in the first period may be profitable, but it is not achievable in a partitional PBE when consumers are patient

[23] In the discrete model, downgrading can never happen as the equilibrium is either fully separating or pooling at the low quality.

[24] By "roughly", we mean to attribute the change in the reservation price induced by the ratchet effect to the second period. Equivalently speaking, $\chi(\cdot)$ is the first-order effect of the static model with the restriction that only one quality is allowed to be offered in the market. See the proof in the Appendix.

($\delta > \frac{1}{\min_{\theta\in[\zeta,\overline{\theta}]} \frac{u_\theta(\phi(\theta),\theta)}{u_{\theta q}(1,\theta)}}$) since in equilibrium the ratchet effect makes all consumers take the same contract in the first period.

4 Conclusion

This paper studies a general model of dynamic price discrimination and product line design with customer recognition. The monopolist can discriminate consumers by offering multiple qualities at different prices in each period or by tracking their purchase history and offering different menus of contracts in the second period conditional on consumers' purchasing behavior in the first period. We show that the firm will offer only one quality in the first period when the surplus function is log submodular or consumers are patient, but the first-period quality and market coverage are not necessarily monotone in the discount factor δ. This is also the first paper to point out the phenomenon of downgrading. To fully understand the firm's dynamic strategy of quality provision, however, work remains to be done to solve for the optimal menu of contracts. Here we only consider the case of monopoly. It will be interesting to extend the basic model to competitive environments.

5 Appendix

Proof of Proposition 2.

First we calculate the profit for each strategy. Let p_1^q denote the price of the product with quality q offered by the firm in the first period.

The profit for strategy T_1: Offering high quality to the high type.

In this case, the monopolist can extract all consumer surplus in the second period by offering the high quality product, charging $u(\overline{q},\theta_H)$ to those customers who purchased in the first period, and $u(\overline{q},\theta_L)$ to those customers who didnot make any purchase before. However, a high-type customer may pretend he is of low type by not making any purchase in the first period if the price $p_1^{\overline{q}}$ is too high. Therefore the monopolist's problem is to choose the price $p_1^{\overline{q}}$ to maximize the discounted profit subject to the individual rationality and incentive constraints:

$$
\begin{aligned}
\Pi(T_1) &= \max_{p_1^{\overline{q}}} \ \alpha[p_1^{\overline{q}} - c(\overline{q})] + \delta[\alpha(u(\overline{q},\theta_H) - c(\overline{q})) + (1-\alpha)(u(\overline{q},\theta_L) - c(\overline{q}))] \\
s.t. \ u(\overline{q},\theta_H) - p_1^{\overline{q}} &\geq \delta[u(\overline{q},\theta_H) - u(\overline{q},\theta_L)] \\
u(\overline{q},\theta_L) - p_1^{\overline{q}} &\leq 0
\end{aligned}
$$

Therefore the monopolist will set $p_1^{\overline{q}} = u(\overline{q},\theta_H) - \delta[u(\overline{q},\theta_H) - u(\overline{q},\theta_L)]$ and the profit is:

$$
\begin{aligned}
\Pi(T_1) &= \alpha[u(\overline{q},\theta_H) - \delta(u(\overline{q},\theta_H) - u(\overline{q},\theta_L)) - c(\overline{q})] + \delta[\alpha(u(\overline{q},\theta_H) - c(\overline{q})) + (1-\alpha)(u(\overline{q},\theta_L) - c(\overline{q}))] \\
&= \alpha[u(\overline{q},\theta_H) - c(\overline{q})] + \delta[u(\overline{q},\theta_L) - c(\overline{q})] \\
&= \alpha S(\overline{q},\theta_H) + \delta S(\overline{q},\theta_L)
\end{aligned}
$$

The profit for strategy T_2: Offering high quality to both types.

In this case, the monopolist will simply charge $p_1^{\overline{q}} = u(\overline{q},\theta_L)$ and it gains no information from the first-period purchasing behavior, so the second-period maximization problem is equivalent to the static maximization problem. According to proposition 1, we can write down the profit for this strategy as:

$$\Pi(T_2) = \begin{cases} S(\overline{q}, \theta_L) + \delta[\alpha S(\overline{q}, \theta_H)] & if\ \alpha \geq \max\{\frac{S(\overline{q},\, \theta_L)}{S(\overline{q},\, \theta_H)}, \frac{S(\underline{q},\, \theta_L)}{S(\underline{q},\, \theta_H)}\} \\ S(\overline{q}, \theta_L) + \delta S(\overline{q}, \theta_L) & if\ \alpha \leq \min\{\frac{S(\overline{q},\, \theta_L)}{S(\overline{q},\, \theta_H)}, \frac{S(\overline{q},\, \theta_L) - S(\underline{q},\, \theta_L)}{S(\overline{q},\, \theta_H) - S(\underline{q},\, \theta_H)}\} \\ S(\overline{q}, \theta_L) + \delta[\alpha(S(\underline{q}, \theta_L) + S(\overline{q}, \theta_H) - S(\underline{q}, \theta_H)) \\ +(1-\alpha) S(\underline{q}, \theta_L)] & if\ \frac{S(\overline{q},\, \theta_L) - S(\underline{q},\, \theta_L)}{S(\overline{q},\, \theta_H) - S(\underline{q},\, \theta_H)} \leq \alpha \leq \frac{S(\underline{q},\, \theta_L)}{S(\underline{q},\, \theta_H)} \end{cases}$$

The profit for strategy T_3: Offering two qualities.

Under this strategy, the monopolist learns the types of customers completely, and consequently it extracts all consumer surplus in the second period by offering the high quality product, charging $u(\overline{q}, \theta_H)$ to those customers who purchased $\overline{q}$ in the first period and $u(\overline{q}, \theta_L)$ to the rest. Therefore, the monopolist's problem is to choose the prices $(p_1^{\overline{q}}, p_1^{\underline{q}})$ to maximize the discounted profit subject to the individual rationality and incentive constraints:

$$\begin{aligned} \Pi(T_3) &= \max_{p_1^{\overline{q}}, p_1^{\underline{q}}} \alpha[p_1^{\overline{q}} - c(\overline{q})] + (1-\alpha)[p_1^{\underline{q}} - c(\underline{q})] \\ &\quad +\delta[\alpha(u(\overline{q}, \theta_H) - c(\overline{q})) + (1-\alpha)(u(\overline{q}, \theta_L) - c(\overline{q}))] \\ s.t.\ u(\overline{q}, \theta_H) - p_1^{\overline{q}} &\geq u(\underline{q}, \theta_H) - p_1^{\underline{q}} + \delta[u(\overline{q}, \theta_H) - u(\overline{q}, \theta_L)] \\ u(\underline{q}, \theta_L) - p_1^{\underline{q}} &\geq u(\overline{q}, \theta_L) - p_1^{\overline{q}} \\ u(\underline{q}, \theta_L) - p_1^{\underline{q}} &\geq 0 \end{aligned}$$

Adding two incentive constraints, we get a necessary condition for this strategy to be feasible:

$$u(\overline{q}, \theta_H) + u(\underline{q}, \theta_L) \geq u(\underline{q}, \theta_H) + u(\overline{q}, \theta_L) + \delta[u(\overline{q}, \theta_H) - u(\overline{q}, \theta_L)]$$

or equivalently:

$$\delta \leq 1 - \frac{u(\underline{q}, \theta_H) - u(\underline{q}, \theta_L)}{u(\overline{q}, \theta_H) - u(\overline{q}, \theta_L)} = \delta^*$$

Under the condition $\delta \leq \delta^*$, it is straightforward to solve for the maximizers:

$$\begin{aligned} p_1^{\overline{q}} &= u(\overline{q}, \theta_H) - u(\underline{q}, \theta_H) + u(\underline{q}, \theta_L) - \delta[u(\overline{q}, \theta_H) - u(\overline{q}, \theta_L)] \\ p_1^{\underline{q}} &= u(\underline{q}, \theta_L) \end{aligned}$$

The profit is:

$$\begin{aligned} \Pi(T_3) &= \alpha[u(\overline{q}, \theta_H) - u(\underline{q}, \theta_H) + u(\underline{q}, \theta_L) - \delta[u(\overline{q}, \theta_H) - u(\overline{q}, \theta_L)] - c(\overline{q})] \\ &\quad +(1-\alpha)[u(\underline{q}, \theta_L) - c(\underline{q})] + \delta[\alpha(u(\overline{q}, \theta_H) - c(\overline{q})) + (1-\alpha)(u(\overline{q}, \theta_L) - c(\overline{q}))] \\ &= \alpha[u(\overline{q}, \theta_H) - c(\overline{q}) - (u(\underline{q}, \theta_H) - c(\underline{q}))] + u(\underline{q}, \theta_L) - c(\underline{q}) + \delta[u(\overline{q}, \theta_L) - c(\overline{q})] \\ &= \alpha[S(\overline{q}, \theta_H) - S(\underline{q}, \theta_H)] + S(\underline{q}, \theta_L) + \delta S(\overline{q}, \theta_L) \end{aligned}$$

If $\alpha \geq \max\{\frac{S(\overline{q},\, \theta_L)}{S(\overline{q},\, \theta_H)}, \frac{S(\underline{q},\, \theta_L)}{S(\underline{q},\, \theta_H)}\}$, it can be readily seen that T_1 is the optimal strategy. If $\alpha \leq \min\{\frac{S(\overline{q},\, \theta_L)}{S(\overline{q},\, \theta_H)}, \frac{S(\overline{q},\, \theta_L) - S(\underline{q},\, \theta_L)}{S(\overline{q},\, \theta_H) - S(\underline{q},\, \theta_H)}\}$, T_2 is the optimal strategy. Now suppose $\frac{S(\overline{q},\, \theta_L) - S(\underline{q},\, \theta_L)}{S(\overline{q},\, \theta_H) - S(\underline{q},\, \theta_H)} \leq \alpha \leq$

$\frac{S(\underline{q},\,\theta_L)}{S(\underline{q},\,\theta_H)}$. If $\delta \leq \delta^*$, then T_3 is the optimal strategy. If $\delta > \delta^*$, then T_3 is not feasible, and we need to compare the profits from T_1 and T_2. The profit difference between these two strategies is

$$\begin{aligned} & \Pi(T_1) - \Pi(T_2) \\ = \;& \alpha S(\overline{q}, \theta_H) + \delta S(\overline{q}, \theta_L) - [S(\overline{q}, \theta_L) + \delta[\alpha(S(\underline{q}, \theta_L) + S(\overline{q}, \theta_H) - S(\underline{q}, \theta_H)) + (1-\alpha)S(\underline{q}, \theta_L)]] \\ = \;& (1-\delta)[\alpha S(\overline{q}, \theta_H) - S(\overline{q}, \theta_L)] + \delta[\alpha S(\underline{q}, \theta_H) - S(\underline{q}, \theta_L)] \end{aligned}$$

Hence T_1 is optimal if and only if[25]

$$\delta \leq \delta^c = \frac{\alpha S(\overline{q}, \theta_H) - S(\overline{q}, \theta_L)}{\alpha S(\overline{q}, \theta_H) - S(\overline{q}, \theta_L) - [\alpha S(\underline{q}, \theta_H) - S(\underline{q}, \theta_L)]}.$$

Proof of Proposition 5.

Again, the firm has three strategies in the first period: offering one quality to the high type (T_1), offering one quality to both types (T_2), and offering two qualities (T_3). If the firm is going to offer just one quality, then apparently it will offer $q = 1$ according to Assumption 2. Mimicking the proof of Proposition 2 and applying the result of Proposition 4, the profits from T_1 and T_2 can be easily obtained as[26]

$$\Pi(T_1) = \alpha S(1, \theta_H) + \delta S(1, \theta_L)$$

$$\Pi(T_2) = \begin{cases} S(1,\theta_L) + \delta[\alpha S(1,\theta_H)] & if\;\; \alpha \geq \max\{\frac{S(1,\,\theta_L)}{S(1,\,\theta_H)}, \frac{S(q^*,\,\theta_L)}{S(q^*,\,\theta_H)}\} \\ S(1,\theta_L) + \delta S(1,\theta_L) & if\;\; \alpha \leq \min\{\frac{S(1,\,\theta_L)}{S(1,\,\theta_H)}, \frac{S_q(1,\,\theta_L)}{S_q(1,\,\theta_H)}\} \\ \;S(1,\theta_L) + \delta[\alpha(S(q^*,\theta_L) + S(1,\theta_H) - S(q^*,\theta_H)) & if\;\; \frac{S_q(1,\,\theta_L)}{S_q(1,\,\theta_H)} \leq \alpha \leq \frac{S(q^*,\,\theta_L)}{S(q^*,\,\theta_H)} \\ \;+(1-\alpha)S(q^*,\theta_L)] & \end{cases}$$

The analysis for strategy T_3 is slightly more complicated. If the monopolist adopts strategy T_3, then it will offer the high quality $q_H = 1$ from Assumption 2. Hence, the maximization problem is to choose a low quality $q \in [0, 1)$ and prices (p_1^1, p_1^q) that solve

$$\begin{aligned} \Pi(T_3) \;&=\; \max_{q, p_1^1, p_1^q} \; \alpha[p_1^1 - c(1)] + (1-\alpha)[p_1^q - c(q)] \\ &\quad\; +\delta[\alpha(u(1,\theta_H) - c(1)) + (1-\alpha)(u(1,\theta_L) - c(1))] \\ s.t.\; u(1,\theta_H) - p_1^1 \;&\geq\; u(q,\theta_H) - p_1^q + \delta[u(1,\theta_H) - u(1,\theta_L)] \\ u(q,\theta_L) - p_1^q \;&\geq\; u(1,\theta_L) - p_1^1 \\ u(q,\theta_L) - p_1^q \;&\geq\; 0 \end{aligned}$$

To satisfy the first two inequalities, the following relation must hold

$$\delta \leq \lambda(q) = 1 - \frac{u(q,\theta_H) - u(q,\theta_L)}{u(1,\theta_H) - u(1,\theta_L)}$$

[25] Under the condition $\frac{S(\overline{q},\,\theta_L) - S(\underline{q},\,\theta_L)}{S(\overline{q},\,\theta_H) - S(\underline{q},\,\theta_H)} \leq \alpha \leq \frac{S(\underline{q},\,\theta_L)}{S(\underline{q},\,\theta_H)}$, $\delta^c \in [0, 1]$ only if $\alpha S(\overline{q}, \theta_H) - S(\overline{q}, \theta_L) \geq 0$. Hence T_2 is the optimal strategy directly if $\alpha S(\overline{q}, \theta_H) - S(\overline{q}, \theta_L) < 0$. Moreover, $\delta^c = 1$ when $\alpha S(\underline{q}, \theta_H) = S(\underline{q}, \theta_L)$, and $\delta^c = \delta^*$ when $\alpha = \frac{(1-\delta^*)S(\overline{q},\theta_L) + \delta^* S(\underline{q},\theta_L)}{(1-\delta^*)S(\overline{q},\theta_H) + \delta^* S(\underline{q},\,\theta_H)}$.

[26] Recall that $q^* = \arg\max_{q \in [0,1]} \Gamma(q) = S(q, \theta_L) - \alpha S(q, \theta_H)$.

We notice that $\lambda(q)$ is monotone decreasing in q; hence, it has an inverse function denoted by $\lambda^{-1}(\cdot)$. Let us temporarily ignore this condition and solve it directly. Applying the result in Proposition 2, the profit from T_3 for a given q is

$$\alpha[S(1,\theta_H) - S(q,\theta_H)] + S(q,\theta_L) + \delta S(1,\theta_L),$$

which is maximized at $q = q^*$. However, the relation $\delta \leq \lambda(q^*)$ is not guaranteed to be satisfied. If $\delta > \lambda(q^*)$, then q^* is not feasible anymore, and the firm has to *decrease* the quality to meet this condition. If $\delta > \lambda(0)$, then T_3 is not feasible. If $\delta \leq \lambda(q^*)$, the q^* is the optimal choice. If $\lambda(q^*) \leq \delta \leq \lambda(0)$, then the firm will choose $q = \lambda^{-1}(\delta)$. In sum, the profit for T_3 is

$$\Pi(T_3) = \begin{cases} \alpha[S(1,\theta_H) - S(q^*,\theta_H)] + S(q^*,\theta_L) + \delta S(1,\theta_L) & if\ \delta \leq \lambda(q^*) \\ \alpha[S(1,\theta_H) - S(\lambda^{-1}(\delta),\theta_H)] + S(\lambda^{-1}(\delta),\theta_L) + \delta S(1,\theta_L) & if\ \lambda(q^*) \leq \delta \leq \lambda(0) \\ \text{not feasible} & if\ \delta > \lambda(0) \end{cases}$$

A straightforward comparison gives us Proposition 5. Q.E.D.

Proof of Proposition 8.

Let $\{(p_i, q_i)\}$ be the first-period contracts offered by the firm with $q_i < q_{i+1}$, and $\{\theta_i\}$ be the cutoff points induced by $\{(p_i, q_i)\}$. Obviously $\theta_1 \geq \zeta$, for the firm cannot get any profit in either period if $\theta_1 < \zeta$. Define $A_i = [\theta_i,\ \theta_{i+1}]$. For each A_i the firm chooses in the second period a cutoff point $\theta_{A_i} \in [\theta_i,\ \theta_{i+1}]$ and a nondecreasing function $q(\cdot\ |A_i) : [\theta_{A_i}, \theta_{i+1}] \to [0,1]$ to maximize

$$\int_{\theta_{A_i}}^{\theta_{i+1}} [S(q(\theta|A_i),\theta) - u_\theta(q(\theta|A_i),\theta)\frac{F(\theta_{i+1}) - F(\theta)}{f(\theta)}]f(\theta)d\theta.$$

Let $\{\theta^*_{A_i}, q^*(\cdot\ |A_i)\}$ denote the optimal second-period strategy in A_i. We notice that $q^*(\theta_{i+1}|A_i) = \phi(\theta_{i+1})$. The incentive constraints for θ_{i+1} and $\gamma \in [\theta^*_{A_i}, \theta_{i+1})$ require that

$$\begin{aligned} u(q_{i+1},\theta_{i+1}) - p_{i+1} &\geq u(q_i,\theta_{i+1}) - p_i + \delta \int_{\theta^*_{A_i}}^{\theta_{i+1}} u_\theta(q^*(\theta|A_i),\theta)d\theta \\ u(q_i,\gamma) - p_i + \delta \int_{\theta^*_{A_i}}^{\gamma} u_\theta(q^*(\theta|A_i),\theta)d\theta &\geq u(q_{i+1},\gamma) - p_{i+1} \end{aligned}$$

Therefore for any $\gamma \in [\theta^*_{A_i}, \theta_{i+1})$

$$u(q_{i+1},\theta_{i+1}) - u(q_{i+1},\gamma) - [u(q_i,\theta_{i+1}) - u(q_i,\gamma)] - \delta \int_{\gamma}^{\theta_{i+1}} u_\theta(q^*(\theta|A_i),\theta)d\theta \geq 0.$$

Dividing both sides by $(\theta_{i+1} - \gamma)$ and letting $\gamma \to \theta_{i+1}$ gives us

$$u_\theta(q_{i+1},\theta_{i+1}) - u_\theta(q_i,\theta_{i+1}) \geq \delta u_\theta(\phi(\theta_{i+1}),\theta_{i+1}).$$

By the mean-value theorem, $\exists\ a \in [q_i, q_{i+1}]$ such that

$$u_{\theta q}(a,\theta_{i+1})(q_{i+1} - q_i) = u_\theta(q_{i+1},\theta_{i+1}) - u_\theta(q_i,\theta_{i+1}) \geq \delta u_\theta(\phi(\theta_{i+1}),\theta_{i+1}).$$

Hence[27]

$$q_{i+1} - q_i \geq \delta \frac{u_\theta(\phi(\theta_{i+1}), \theta_{i+1})}{u_{\theta q}(a, \theta_{i+1})} \geq \delta \frac{u_\theta(\phi(\theta_{i+1}), \theta_{i+1})}{u_{\theta q}(1, \theta_{i+1})} \geq \delta \min_{\theta \in [\zeta, \overline{\theta}]} \frac{u_\theta(\phi(\theta), \theta)}{u_{\theta q}(1, \theta)}.$$

Q.E.D.

Proof of Proposition 9.

Let (p_1, q_1) be the first-period contract offered by the firm and θ_{1l} be the first-period cutoff point induced by (p_1, q_1). Define $A = [\theta_{1l}, \overline{\theta}]$ and $N = [\underline{\theta}, \theta_{1l}]$. In the second period, the firm can offer different menus of contracts in A and N. Let θ_{Al} be the cutoff point in A, θ_{Nl} the cutoff point in N, $q_2(\cdot\ |A) : [\theta_{Al}, \overline{\theta}] \to [0, 1]$ a nondecreasing quality offering in A, and $q_2(\cdot\ |N) : [\theta_{Nl}, \overline{\theta}] \to [0, 1]$ a nondecreasing quality offering in N. Then the firm's profit function is

$$\begin{aligned}
\Pi &= [1 - F(\theta_{1l})][p_1 - c(q_1)] + \delta\{\int_{\theta_{Al}}^{\overline{\theta}} [u(q_2(\theta|A), \theta) - c(q_2(\theta|A)) - u_\theta(q_2(\theta|A), \theta)\frac{1 - F(\theta)}{f(\theta)}]f(\theta)d\theta \\
&\quad + \int_{\theta_{Nl}}^{\theta_{1l}} [u(q_2(\theta|N), \theta) - c(q_2(\theta|N)) - u_\theta(q_2(\theta|N), \theta)\frac{F(\theta_{1l}) - F(\theta)}{f(\theta)}]f(\theta)d\theta\} \\
&= [1 - F(\theta_{1l})][p_1 - c(q_1)] + \delta\{\int_{\theta_{Al}}^{\overline{\theta}} [S(q_2(\theta|A), \theta) - u_\theta(q_2(\theta|A), \theta)\frac{1 - F(\theta)}{f(\theta)}]f(\theta)d\theta \\
&\quad + \int_{\theta_{Nl}}^{\theta_{1l}} [S(q_2(\theta|N), \theta) - u_\theta(q_2(\theta|N), \theta)\frac{F(\theta_{1l}) - F(\theta)}{f(\theta)}]f(\theta)d\theta\}
\end{aligned}$$

To solve this problem, let us determine the first-period price p_1. Consider consumer $\theta_{1l}'s$ behavior. In the first period, θ_{1l} can choose to be the lowest type in A and gain no surplus in the second period, or the highest type in N and acquire some information rent. To satisfy the incentive constraint, p_1 should be set to make θ_{1l} indifferent between these two alternatives. By taking the contract (p_1, q_1), consumer $\theta_{1l}'s$ utility is $u(q_1, \theta_{1l}) - p_1$. If he rejects the contract, then his utility will be $\delta \int_{\theta_{Nl}}^{\theta_{1l}} u_\theta(q_2(\theta|N), \theta)d\theta$ by the envelope theorem. Hence in equilibrium it must be the case that $u(q_1, \theta_{1l}) - p_1 = \delta \int_{\theta_{Nl}}^{\theta_{1l}} u_\theta(q_2(\theta|N), \theta)d\theta$, or equivalently $p_1 = u(q_1, \theta_{1l}) - \delta \int_{\theta_{Nl}}^{\theta_{1l}} u_\theta(q_2(\theta|N), \theta)d\theta$. Then the profit function can be rewritten as

$$\begin{aligned}
\Pi &= [1 - F(\theta_{1l})][u(q_1, \theta_{1l}) - c(q_1) - \delta \int_{\theta_{Nl}}^{\theta_{1l}} u_\theta(q_2(\theta|N), \theta)d\theta] \\
&\quad + \delta\{\int_{\theta_{Al}}^{\overline{\theta}} [S(q_2(\theta|A), \theta) - u_\theta(q_2(\theta|A), \theta)\frac{1 - F(\theta)}{f(\theta)}]f(\theta)d\theta \\
&\quad + \int_{\theta_{Nl}}^{\theta_{1l}} [S(q_2(\theta|N), \theta) - u_\theta(q_2(\theta|N), \theta)\frac{F(\theta_{1l}) - F(\theta)}{f(\theta)}]f(\theta)d\theta\}
\end{aligned}$$

The optimal choice of q_1 must solve $\max_{q_1 \in [0,1]} u(q_1, \theta_{1l}) - c(q_1)$. Hence $q_1 = \phi(\theta_{1l})$. On the other hand, we notice that the optimal choice of $q_2(\theta_{1l}|N)$ is $\phi(\theta_{1l})$ as well. Furthermore, we also need to

[27] Recall that $S_{qq\theta}(q, \theta) \geq 0$.

check the incentive compatibility for any $\theta \in A$ and $\theta \in N$. Given that $q_1 = q_2(\theta_{1l}|N) = \phi(\theta_{1l})$, it is straightforward to verify that this is indeed the case.

For a given θ_{1l}, let $\{\theta^*_{Al}, \theta^*_{Nl}, q^*_2(\cdot\,|A), q^*_2(\cdot\,|N)\}$ denote the optimal second-period strategy. To determine the optimal value of θ_{1l}, we analyze the first-order condition for θ_{1l}.[28] We claim that $\theta_{1l} > \theta^*_l$ by showing that $\frac{\partial \Pi}{\partial \theta_{1l}} > 0\ \forall \theta \le \theta^*_l$.[29] There are two cases to be considered: $\theta^*_{Al} = \theta_{1l}$ and $\theta^*_{Al} > \theta_{1l}$.

Case 1. $\theta^*_{Al} > \theta_{1l}$. Using the envelope theorem, $\frac{\partial \Pi}{\partial \theta_{1l}}$ in this case is :

$$\begin{aligned}
\frac{\partial \Pi}{\partial \theta_{1l}} &= -f(\theta_{1l})[S(q_1,\theta_{1l}) - \delta \int_{\theta^*_{Nl}}^{\theta_{1l}} u_\theta(q^*_2(\theta|N),\theta)d\theta] + [1-F(\theta_{1l})][u_\theta(q_1,\theta_{1l}) \\
&\quad - \frac{\partial}{\partial \theta_{1l}} \delta \int_{\theta^*_{Nl}}^{\theta_{1l}} u_\theta(q^*_2(\theta|N),\theta)d\theta] + \delta[S(q^*_2(\theta_{1l}|N),\theta_{1l})f(\theta_{1l}) - \int_{\theta^*_{Nl}}^{\theta_{1l}} u_\theta(q^*_2(\theta|N),\theta)d\theta f(\theta_{1l})] \\
&= [(1-F(\theta_{1l}))u_\theta(q_1,\theta_{1l}) - f(\theta_{1l})S(q_1,\theta_{1l})] \\
&\quad + \delta[S(q^*_2(\theta_{1l}|N),\theta_{1l})f(\theta_{1l}) - (1-F(\theta_{1l}))\frac{\partial}{\partial \theta_{1l}} \int_{\theta^*_{Nl}}^{\theta_{1l}} u_\theta(q^*_2(\theta|N),\theta)d\theta] \\
&= [(1-F(\theta_{1l}))u_\theta(q_1,\theta_{1l}) - f(\theta_{1l})S(q_1,\theta_{1l})] + \delta[S(q^*_2(\theta_{1l}|N),\theta_{1l})f(\theta_{1l}) - (1-F(\theta_{1l}))u_\theta(q^*_2(\theta_{1l}|N),\theta_{1l})] \\
&\quad - \delta(1-F(\theta_{1l}))[\int_{\theta^*_{Nl}}^{\theta_{1l}} \frac{\partial}{\partial \theta_{1l}} u_\theta(q^*_2(\theta|N),\theta)d\theta - u_\theta(q^*_2(\theta^*_{Nl}|N),\theta^*_{Nl})\frac{\partial \theta^*_{Nl}}{\partial \theta_{1l}}] \\
&= -(1-\delta)f(\theta_{1l})[S(\phi(\theta_{1l}),\theta_{1l}) - \frac{1-F(\theta_{1l})}{f(\theta_{1l})}u_\theta(\phi(\theta_{1l}),\theta_{1l})] \\
&\quad - \delta(1-F(\theta_{1l}))[\int_{\theta^*_{Nl}}^{\theta_{1l}} \frac{\partial}{\partial \theta_{1l}} u_\theta(q^*_2(\theta|N),\theta)d\theta - u_\theta(q^*_2(\theta^*_{Nl}|N),\theta^*_{Nl})\frac{\partial \theta^*_{Nl}}{\partial \theta_{1l}}]
\end{aligned}$$

Fact 1. $\frac{\partial}{\partial \theta_{1l}} u_\theta(q^*_2(\theta|N),\theta) \le 0$ and $\frac{\partial \theta^*_{Nl}}{\partial \theta_{1l}} > 0$. Invoking the envelope theorem on the second-period value function for some θ on N,

$$\Psi(\theta_{1l},\theta) = \max_{q_2(\theta|N)\in[0,1]} \Phi(q_2(\theta|N)) = S(q_2(\theta|N),\theta) - u_\theta(q_2(\theta|N),\theta)\frac{F(\theta_{1l}) - F(\theta)}{f(\theta)},$$

we get $\frac{\partial}{\partial \theta_{1l}}\Psi(\theta_{1l},\theta) = -u_\theta(q^*_2(\theta|N),\theta)f(\theta_{1l})/f(\theta) < 0$; hence, θ^*_{Nl} is increasing in θ_{1l}. $q^*_2(\theta|N)$ solves the first order condition $\Phi_q = 0$ (or $\Phi_q > 0$ if $q^*_2(\theta|N) = 1$). Applying the implicit function theorem, we can get $\frac{\partial}{\partial \theta_{1l}} q^*_2(\theta|N) \le 0$. Thus

$$\frac{\partial}{\partial \theta_{1l}} u_\theta(q^*_2(\theta|N),\theta) = u_{\theta q}(q^*_2(\theta|N),\theta)\frac{\partial}{\partial \theta_{1l}} q^*_2(\theta|N) \le 0.$$

Fact 2. If $\theta_{1l} \le \theta^*_l$, then we have

$$S(\phi(\theta_{1l}),\theta_{1l}) - \frac{1-F(\theta_{1l})}{f(\theta_{1l})}u_\theta(\phi(\theta_{1l}),\theta_{1l}) \le 0,$$

[28] Here we simply assume that the second- order condition is satisfied.

[29] Recall that θ^*_l is the optimal cutoff point in the static model.

for $S(q,\theta_l^*) - \frac{1-F(\theta_l^*)}{f(\theta_l^*)} u_\theta(q,\theta_l^*)\ \forall q \in [0,1]$ is zero at most and $\theta_{1l} \le \theta_l^*$.

Combining these two facts, we conclude that $\frac{\partial \Pi}{\partial \theta_{1l}} > 0$ when $\theta_{1l} \le \theta_l^*$. Hence, the optimal choice $\theta_{1l}^* \in (\theta_l^*, \overline{\theta})$ in this case.

Case 2. Now suppose $\theta_{Al}^* = \theta_{1l}$. Then $\frac{\partial \Pi}{\partial \theta_{1l}}$ has one more term:

$$-\delta[S(q_2(\theta_{1l}|A),\theta_{1l}) - u_\theta(q_2(\theta_{1l}|A),\theta_{1l})\frac{1-F(\theta_{1l})}{f(\theta_{1l})}]f(\theta_{1l}),$$

which is non-negative when $\theta_{1l} \le \theta_l^*$. Therefore, the optimal choice $\theta_{1l}^* \in (\theta_l^*, \overline{\theta})$.

Next we study the comparative statics. Since $\theta_{1l}^* \in (\theta_l^*, \overline{\theta})$, it is sufficient for us to analyze $\frac{\partial \Pi}{\partial \theta_{1l}}$ for $\theta_{1l} \in (\theta_l^*, \overline{\theta})$. In this range, $\theta_{Al} = \theta_{1l}$, and $\frac{\partial \Pi}{\partial \theta_{1l}}$ can be expressed as $\frac{\partial \Pi}{\partial \theta_{1l}} = -f(\theta_{1l})\chi(\theta_{1l}) + \delta\gamma(\theta_{1l})$, where

$$\begin{aligned}
\chi(\theta_{1l}) &= S(\phi(\theta_{1l}),\theta_{1l}) - \frac{1-F(\theta_{1l})}{f(\theta_{1l})} u_\theta(\phi(\theta_{1l}),\theta_{1l}) \\
\gamma(\theta_{1l}) &= f(\theta_{1l})[\chi(\theta_{1l}) - [S(q_2(\theta_{1l}|A),\theta_{1l}) - u_\theta(q_2(\theta_{1l}|A),\theta_{1l})\frac{1-F(\theta_{1l})}{f(\theta_{1l})}]] \\
&\quad -(1-F(\theta_{1l}))[\int_{\theta_{Nl}^*}^{\theta_{1l}} \frac{\partial}{\partial \theta_{1l}} u_\theta(q_2^*(\theta|N),\theta)d\theta - u_\theta(q_2^*(\theta_{Nl}^*|N),\theta_{Nl}^*)\frac{\partial \theta_{Nl}^*}{\partial \theta_{1l}}]
\end{aligned}$$

Therefore we have $\frac{\partial^2 \Pi}{\partial \theta_{1l} \partial \delta} = \gamma(\theta_{1l})$. If $\gamma(\theta_{1l}^*) > 0$, then Π is (locally) supermodular at (δ, θ_{1l}^*), which implies that the optimal first-period cutoff point θ_{1l}^* is increasing in δ. Similarly, if $\gamma(\theta_{1l}^*) < 0$, then θ_{1l}^* is decreasing in δ. So we have *sign* $\frac{\partial \theta_{1l}^*}{\partial \delta} =$ *sign* $\gamma(\theta_{1l}^*)$. The first-order condition $\frac{\partial \Pi}{\partial \theta_{1l}}|_{\theta_{1l}=\theta_{1l}^*} = 0$ implies that *sign* $\chi(\theta_{1l}^*) =$ *sign* $\gamma(\theta_{1l}^*)$. Thus we have *sign* $\frac{\partial \theta_{1l}^*}{\partial \delta} =$ *sign* $\chi(\theta_{1l}^*)$. On the other hand, we observe that $S(q,\theta)$ is supermodular on (q,θ) by the single crossing property. Hence the optimal first-period quality $q_1^* = \phi(\theta)$, which maximizes the social surplus at θ, is increasing in θ. Therefore *sign* $\frac{\partial q_1^*}{\partial \delta} =$ *sign* $\frac{\partial \theta_{1l}^*}{\partial \delta}$. Q.E.D.

Proof of Proposition 10.

Case 1. When δ is high. This is a direct consequence of proposition 8 and the fact that $q \in [0,1]$.

Case 2. When $S(q,\theta)$ is log submodular. First, we show that any two-contract strategy is strictly dominated by a single-contract strategy when $S(q,\theta)$ is log submodular. To make the analysis clear, we divide the proof into three steps.

Step 1. Write down the profit function. Let $C_1 = \{(p_i,q_i)\}_{i=1}^2$ be any two-contract strategy with $q_1 < q_2$, and $\{\theta_i\}_{i=1}^2$ with $\theta_1 < \theta_2$ the cutoff points induced by C_1. Define $A_i = [\theta_i, \theta_{i+1}]$ (let $\theta_3 = \overline{\theta}$) and $N = [\underline{\theta}, \theta_1]$. Let $\{\theta_{A_i}^*, q^*(\cdot\,|A_i)\}$ denote the optimal strategy in A_i, and $\{\theta_N^*, q^*(\cdot\,|N)\}$ the optimal strategy in N. As explained in the proof of proposition 9, (p_1, p_2) must satisfy the following equations:

$$\begin{aligned}
u(q_1,\theta_1) - p_1 &= \delta \int_{\theta_N^*}^{\theta_1} u_\theta(q^*(\theta|N),\theta)d\theta \\
u(q_2,\theta_2) - p_2 &= u(q_1,\theta_2) - p_1 + \delta \int_{\theta_{A_1}^*}^{\theta_2} u_\theta(q^*(\theta|A_1),\theta)d\theta
\end{aligned}$$

Thus the profit function can be written as:

$$\begin{aligned}
\Pi &= [1-F(\theta_2)][u(q_2,\theta_2)-u(q_1,\theta_2)+u(q_1,\theta_1)-\delta\int_{\theta_N^*}^{\theta_1} u_\theta(q^*(\theta|N),\theta)d\theta-\delta\int_{\theta_{A_1}^*}^{\theta_2} u_\theta(q^*(\theta|A_1),\theta)d\theta-c(q_2)] \\
&\quad +[F(\theta_2)-F(\theta_1)][u(q_1,\theta_1)-\delta\int_{\theta_N^*}^{\theta_1} u_\theta(q^*(\theta|N),\theta)d\theta-c(q_1)] \\
&\quad +\delta\{\int_{\theta_{A_2}^*}^{\overline{\theta}} [S(q^*(\theta|A_2),\theta)-u_\theta(q^*(\theta|A_2),\theta)\frac{1-F(\theta)}{f(\theta)}]f(\theta)d\theta \\
&\quad +\int_{\theta_{A_1}^*}^{\theta_2} [S(q^*(\theta|A_1),\theta)-u_\theta(q^*(\theta|A_1),\theta)\frac{F(\theta_2)-F(\theta)}{f(\theta)}]f(\theta)d\theta \\
&\quad +\int_{\theta_N^*}^{\theta_1} [S(q^*(\theta|N),\theta)-u_\theta(q^*(\theta|N),\theta)\frac{F(\theta_1)-F(\theta)}{f(\theta)}]f(\theta)d\theta\}
\end{aligned}$$

Step 2. We claim that $\theta_1 > \theta_l^*$. There are two cases to be considered: $\theta_{A_1}^* = \theta_1$ and $\theta_{A_1}^* > \theta_1$. First, let us assume $\theta_{Al}^* > \theta_{1l}$. Using the envelope theorem, the first order condition w.r.t. θ_1 in this case is :

$$\begin{aligned}
\frac{\partial\Pi}{\partial\theta_1} &= [1-F(\theta_2)][u_\theta(q_1,\theta_1)-\delta\frac{\partial}{\partial\theta_1}\int_{\theta_N^*}^{\theta_1} u_\theta(q^*(\theta|N),\theta)d\theta] \\
&\quad -f(\theta_1)[u(q_1,\theta_1)-\delta\int_{\theta_N^*}^{\theta_1} u_\theta(q^*(\theta|N),\theta)d\theta-c(q_1)] \\
&\quad +[F(\theta_2)-F(\theta_1)][u_\theta(q_1,\theta_1)-\delta\frac{\partial}{\partial\theta_1}\int_{\theta_N^*}^{\theta_1} u_\theta(q^*(\theta|N),\theta)d\theta] \\
&\quad +\delta[S(q^*(\theta_1|N),\theta_1)f(\theta_1)-f(\theta_1)\int_{\theta_N^*}^{\theta_1} u_\theta(q^*(\theta|N),\theta)d\theta] \\
&= [1-F(\theta_1)][u_\theta(q_1,\theta_1)-\delta\frac{\partial}{\partial\theta_1}\int_{\theta_N^*}^{\theta_1} u_\theta(q^*(\theta|N),\theta)d\theta]-f(\theta_1)S(q_1,\theta_1)+\delta S(q^*(\theta_1|N),\theta_1)f(\theta_1) \\
&= [(1-F(\theta_1))u_\theta(q_1,\theta_1)-f(\theta_1)S(q_1,\theta_1)]+\delta[f(\theta_1)S(q^*(\theta_1|N),\theta_1)-(1-F(\theta_1))u_\theta(q^*(\theta_1|N),\theta_1)] \\
&\quad -\delta(1-F(\theta_1))[\int_{\theta_N^*}^{\theta_1}\frac{\partial}{\partial\theta_1}u_\theta(q^*(\theta|N),\theta)d\theta-u_\theta(q^*(\theta_N^*|N),\theta_N^*)\frac{\partial\theta_N^*}{\partial\theta_1}]
\end{aligned}$$

Fact 1. $\int_{\theta_N^*}^{\theta_1}\frac{\partial}{\partial\theta_1}u_\theta(q^*(\theta|N),\theta)d\theta-u_\theta(q^*(\theta_N^*|N),\theta_N^*)\frac{\partial\theta_N^*}{\partial\theta_1} < 0$, as demonstrated in the proof of proposition 9.

Fact 2. $q^*(\theta_1|N)=1$, since $S(q,\theta)$ is log submodular.

Fact 3. $\delta \le \frac{u_\theta(q_1,\theta_1)}{u_\theta(q^*(\theta_1|N),\theta_1)}$ from the incentive constraint.

Fact 4. $f(\theta_1)S(q^*(\theta_1|N),\theta_1)-(1-F(\theta_1))u_\theta(q^*(\theta_1|N),\theta_1) \le 0$ if $\theta_1 \le \theta_l^*$, since $S(q,\theta_l^*)-\frac{1-F(\theta_l^*)}{f(\theta_l^*)}u_\theta(q,\theta_l^*)$ is zero at most and $\theta_1 \le \theta_l^*$.

Fact 5. $(1-F(\theta_1))u_\theta(q_1,\theta_1)-f(\theta_1)S(q_1,\theta_1) \ge 0$ if $\theta_1 \le \theta_l^*$, since $S(q,\theta_l^*)-\frac{1-F(\theta_l^*)}{f(\theta_l^*)}u_\theta(q,\theta_l^*)$ is zero at most and $\theta_1 \le \theta_l^*$.

Fact 6. We show that if $\theta_1 \leq \theta_l^*$, then

$$[(1-F(\theta_1))u_\theta(q_1,\theta_1) - f(\theta_1)S(q_1,\theta_1)] + \delta[f(\theta_1)S(q^*(\theta_1|N),\theta_1) - (1-F(\theta_1))u_\theta(q^*(\theta_1|N),\theta_1)] > 0.$$

From Facts 2-5, we can get

$$\begin{aligned}
& [(1-F(\theta_1))u_\theta(q_1,\theta_1) - f(\theta_1)S(q_1,\theta_1)] + \delta[f(\theta_1)S(q^*(\theta_1|N),\theta_1) - (1-F(\theta_1))u_\theta(q^*(\theta_1|N),\theta_1)] \\
= \; & [(1-F(\theta_1))u_\theta(q_1,\theta_1) - f(\theta_1)S(q_1,\theta_1)] + \delta[f(\theta_1)S(1,\theta_1) - (1-F(\theta_1))u_\theta(1,\theta_1)] \\
\geq \; & [(1-F(\theta_1))u_\theta(q_1,\theta_1) - f(\theta_1)S(q_1,\theta_1)] + \frac{u_\theta(q_1,\theta_1)}{u_\theta(1,\theta_1)}[f(\theta_1)S(1,\theta_1) - (1-F(\theta_1))u_\theta(1,\theta_1)] \\
= \; & \frac{f(\theta_1)}{u_\theta(1,\theta_1)}[u_\theta(q_1,\theta_1)S(1,\theta_1) - u_\theta(1,\theta_1)S(q_1,\theta_1)] \\
= \; & \frac{f(\theta_1)S(1,\theta_1)S(q_1,\theta_1)}{u_\theta(1,\theta_1)}[\frac{u_\theta(q_1,\theta_1)}{S(q_1,\theta_1)} - \frac{u_\theta(1,\theta_1)}{S(1,\theta_1)}] > 0.
\end{aligned}$$

The last inequality comes from the log submodularity of $S(q,\theta)$.

Combining Facts 1 and 6, we conclude that $\frac{\partial \Pi}{\partial \theta_1} > 0$ for any $\theta_1 \leq \theta_l^*$. Hence the optimal choice of θ_1 is strictly greater than θ_l^*.

Now consider the second case $\theta_{A_1}^* = \theta_1$. In this case, the change of θ_1 has two additional effects. The first order condition w.r.t. θ_1 becomes:

$$\begin{aligned}
\frac{\partial \Pi}{\partial \theta_1} = \; & [(1-F(\theta_1))u_\theta(q_1,\theta_1) - f(\theta_1)S(q_1,\theta_1)] + \delta[f(\theta_1)S(q^*(\theta_1|N),\theta_1) - (1-F(\theta_1))u_\theta(q^*(\theta_1|N),\theta_1)] \\
& + \delta(1-F(\theta_2))u_\theta(q^*(\theta_1|A_1),\theta_1) - \delta[f(\theta_1)S(q^*(\theta_1|A_1),\theta_1) - (F(\theta_2)-F(\theta_1))u_\theta(q^*(\theta_1|A_1),\theta_1)] \\
& - \delta(1-F(\theta_1))[\int_{\theta_N^*}^{\theta_1} \frac{\partial}{\partial \theta_1} u_\theta(q^*(\theta|N),\theta)d\theta - u_\theta(q^*(\theta_N^*|N),\theta_N^*)\frac{\partial \theta_N^*}{\partial \theta_1}] \\
= \; & [(1-F(\theta_1))u_\theta(q_1,\theta_1) - f(\theta_1)S(q_1,\theta_1)] + \delta[f(\theta_1)S(q^*(\theta_1|N),\theta_1) - (1-F(\theta_1))u_\theta(q^*(\theta_1|N),\theta_1)] \\
& - \delta[f(\theta_1)S(q^*(\theta_1|A_1),\theta_1) - (1-F(\theta_1))u_\theta(q^*(\theta_1|A_1),\theta_1)] \\
& - \delta(1-F(\theta_1))[\int_{\theta_N^*}^{\theta_1} \frac{\partial}{\partial \theta_1} u_\theta(q^*(\theta|N),\theta)d\theta - u_\theta(q^*(\theta_N^*|N),\theta_N^*)\frac{\partial \theta_N^*}{\partial \theta_1}]
\end{aligned}$$

The log submodularity of $S(q,\theta)$ implies that $q^*(\theta_1|A_1) = q^*(\theta_1|N) = 1$, therefore

$$\begin{aligned}
\frac{\partial \Pi}{\partial \theta_1} = \; & [(1-F(\theta_1))u_\theta(q_1,\theta_1) - f(\theta_1)S(q_1,\theta_1)] \\
& - \delta(1-F(\theta_1))[\int_{\theta_N^*}^{\theta_1} \frac{\partial}{\partial \theta_1} u_\theta(q^*(\theta|N),\theta)d\theta - u_\theta(q^*(\theta_N^*|N),\theta_N^*)\frac{\partial \theta_N^*}{\partial \theta_1}].
\end{aligned}$$

As shown in the first case, this term is positive if $\theta_1 \leq \theta_l^*$. Hence $\theta_1 > \theta_l^*$.

Step 3. We show that any two-contract strategy is dominated by a one-contract strategy. From step 2, we know that $\theta_1 > \theta_l^*$. It is then easy to see that $\theta_{A_2}^* = \theta_2$ and $\theta_{A_1}^* = \theta_1$. We also know that $q^*(\theta|N) = q^*(\theta|A_1) = q^*(\theta|A_2) = 1$; hence, the profit of this two-contract strategy is

$$\Pi = [1-F(\theta_2)][u(q_2,\theta_2) - u(q_1,\theta_2) + u(q_1,\theta_1) - \delta \int_{\theta_N^*}^{\theta_2} u_\theta(1,\theta)d\theta - c(q_2)]$$

$$+[F(\theta_2)-F(\theta_1)][S(q_1,\theta_1)-\delta\int_{\theta_N^*}^{\theta_1}u_\theta(1,\theta)d\theta]+\delta\{\int_{\theta_1}^{\overline{\theta}}S(1,\theta)f(\theta)d\theta-\int_{\theta_2}^{\overline{\theta}}u_\theta(1,\theta)(1-F(\theta))d\theta$$
$$-\int_{\theta_1}^{\theta_2}u_\theta(1,\theta)(F(\theta_2)-F(\theta))d\theta+\int_{\theta_N^*}^{\theta_1}[S(1,\theta)-u_\theta(1,\theta)\frac{F(\theta_1)-F(\theta)}{f(\theta)}]f(\theta)d\theta\}$$

Now if we just offer one quality q_2 and set the cutoff point at θ_1, then the profit is:

$$\begin{aligned}\widetilde{\Pi} &= (1-F(\theta_1))[u(q_2,\theta_1)-\delta\int_{\theta_N^*}^{\theta_1}u_\theta(1,\theta)d\theta-c(q_2)]\\ &\quad+\delta\{\int_{\theta_1}^{\overline{\theta}}[S(1,\theta)-u_\theta(1,\theta)\frac{1-F(\theta)}{f(\theta)}]f(\theta)d\theta+\int_{\theta_N^*}^{\theta_1}[S(1,\theta)-u_\theta(1,\theta)\frac{F(\theta_1)-F(\theta)}{f(\theta)}]f(\theta)d\theta\}\end{aligned}$$

After rearranging terms, the difference is:

$$\begin{aligned}\widetilde{\Pi}-\Pi &= [1-F(\theta_2)]S(q_1,\theta_2)-[1-F(\theta_1)]S(q_1,\theta_1)+[1-F(\theta_1)]S(q_2,\theta_1)-[1-F(\theta_2)]S(q_2,\theta_2)\\ &= [1-F(\theta_1)]\int_{q_1}^{q_2}S_q(q,\theta_1)dq-[1-F(\theta_2)]\int_{q_1}^{q_2}S_q(q,\theta_2)dq\\ &= \int_{q_1}^{q_2}\int_{\theta_1}^{\theta_2}[S_q(q,\theta)f(\theta)-(1-F(\theta))S_{q\theta}(q,\theta)]d\theta dq>0\end{aligned}$$

by the log submodularity of $S(q,\theta)$. Therefore the firm will not offer two contracts in the first period.

The N-contract case has a recursive structure, and we can duplicate the procedure of the proof above to show that there is a (N-1)-contract that dominates this N-contract strategy. Repeating this argument, we conclude that the firm's optimal strategy is to offer a single contract in the first period. Q.E.D.

References

[1] Acquisti, A., and H. R. Varian, "Conditioning Prices on Purchase History," *Marketing Science*, 24 (2005), 367-381.

[2] Anderson, E. T., and I. Song, "Coordinating Price Reductions and Coupon Events," working paper (2004), Northwestern University.

[3] Anderson, E. T., and J. Dana, "When Is Price Discrimination Profitable," working paper (2006).

[4] Armstrong, M., "Economics Models of Price Discrimination," working paper (2006), University College London.

[5] Bester, H., and R. Strausz, "Contracting with Imperfect Commitment and The Revelation Principle: The Single Agent Case," *Econometrica*, 69 (2001), 1077-1098.

[6] Bhargava, H. K. and V. Choudhary, "Information Goods and Vertical Differentiation," *Journal of Management Information Systems*, 18(2) (2001), 85-102.

[7] Binkley, A., "Casino Chain Mines Data on Gamblers, and Strikes Pay Dirt with Low-rollers," *Wall Street Journal*, May 4, 2000.

[8] Chen, Yongmin, "Paying Customers to Switch," *Journal of Economics and Management Strategy*, 6 (1997), 877-897.

[9] Courty, P. and H. Li, "Sequential Screening," *Review of Economic Studies*, 67 (2000), 697-717.

[10] Deneckere, R. J. and R. P. McAfee, "Damaged Goods," *Journal of Economics and Management Strategy*, 5(2) (1996), 149-174.

[11] Fudenberg, D. and J. Tirole, "Upgrades, Trade-ins, and Buybacks," *Rand Journal of Economics*, 29 (1998), 238-258.

[12] Fudenberg, D. and J. Tirole, "Customer Poaching and Brand Switching," *Rand Journal of Economics*, 31 (2000), 634-657.

[13] Fudenberg, D. and J. M. Villas-Boas, "Behavior-Based Price Discrimination and Customer Recognition," *Economics and Information Systems* (2007), Elsevier, Oxford.

[14] Gale, I. L. and Holmes, T. J., "Advance-Purchase Discounts and Monopoly Allocation of Capacity," *American Economic Review*, 83(1) (1993), 135-146.

[15] Hart, O. D. and J. Tirole, "Contract Renegotiation and Coasian Dynamics," *Review of Economic Studies*, 55 (1988), 509-540.

[16] Johnson, Justin and David Myatt, "Multiproduct Quality Competition: Fighting Brands and Product Line Pruning," *American Economic Review*, 93(3) (2003), 748-774.

[17] Laffont, J. J. and J. Tirole, "The Dynamics of Incentive Contracts," *Econometrica*, 56 (1988), 1153-1175.

[18] McAfee, Preston, "Damaged Goods Pricing," working paper (2006), California Institute of Technology.

[19] Milgrom, P. and I. Segal, "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70 (2002), 583-601.

[20] Mussa, M. and S. Rosen, "Monopoly and Product Quality," *Journal of Economic Theory*, 18(2) (1978), 301-317.

[21] Nevo, A. and C. Wolfram, "Why Do Manufacturers Issue Coupons? An Empirical Analysis of Breakfast Cereals," *Rand Journal of Economics*, 33(2) (2002), 319-339.

[22] Salant, S. W., "When is Inducing Self-Selection Suboptimal for a Monopolist?," *Quarterly Journal of Economics*, 104(2) (1989), 391-397.

[23] Stokey, N. L., "Intertemporal Price Discrimination," *Quarterly Journal of Economics*, 93(3) (1979), 355-371.

[24] Stole, L., "Price Discrimination in Competitive Environments," *Handbook of Industrial Organization*, 3 (2005).

[25] Streitfield, D. "On the Web Price Tags Blur: What You Pay Could Depend on Who You Are," *The Washington Post*, September 27, 2000, P. A1.

[26] Taylor, C. R., "Supplier Surfing: Price-Discrimination in Markets with Repeat Purchases," *Rand Journal of Economics*, 34 (2003), 223-246.

[27] Taylor, C. R., "Consumer Privacy and The Market for Customer Information," *Rand Journal of Economics*, 35 (2004), 631-651.

[28] Turow, J., Feldman L. and K. Meltzer, "Open to Exploitation: American Shoppers Online and Offline," Annenberg Public Policy Center, University of Pennsylvania.

[29] Varian, H. R., "Versioning information goods," in *Internet Publishing and Beyond: The Economics of Digital Information and Intellectual Property*, ed. by B. Kahin and H. R. Varian. MIT Press (2001).

[30] Villas-Boas, J. M., "Product Line Design for a Distribution Channel," *Marketing Science*, 17(2) (1998), 604-631.

[31] Villas-Boas, J. M., "Dynamic Competition with Customer Recognition," *Rand Journal of Economics*, 30 (1999), 604-631.

[32] Villas-Boas, J. M., "Price Cycles in Markets with Customer Recognition," *Rand Journal of Economics*, 35 (2004), 486-501.

[33] Weiss, R. M. and A. K. Mehrotra, "Online Dynamic Pricing: Efficiency, Equity and the Future of E-commerce," *Virginia Journal of Law and Technology*, 6(2) (2001), A 11.

[34] Weitzman, O., "The 'Ratchet Principle' and Performance Incentives," *Bell Journal of Economics*, 11 (1980), 302-308.

[35] Zhang, J., "Dynamic Segmentation and Product Design," working paper (2005), University of California, Berkeley.