Consumer Learning and Price Discrimination*

Jaesoo Kim[†] Stephen Salant[‡] Dongsoo Shin[§]

2012

Very Preliminary

*We thank Ching-to Albert Ma, Subir Chakrabarti, Henry Mak, and Huanxing Yang for helpful discussions.

[†]Department of Economics, Indiana University-Purdue University Indianapolis, Indianapolis, IN 46202, **E-mail:** jaeskim@iupui.edu

[‡]Department of Economics, University of Michigan, Ann Arbor, MI 48109, **E-mail:** ssalant@umich.edu

[§]Department of Economics, Leavey School of Business, Santa Clara University, Santa Clara, CA 95053, **E-mail:** dshin@scu.edu

Consumer Learning and Price Discrimination

ABSTRACT

We study the monopolistic firm's selling mechanism when consumers are required to learn how to use products. We compare the case where the firm helps consumers learn the product features and benefits, to the case where the firm does not provide any help. In the latter case, the firm may design the products to induce self-learning by a subset of consumers. A consumer's preference is private information. We show that, for small or large learning costs, it is optimal for the firm to directly incur the learning cost. For intermediate learning costs, however, making the consumers to incur the learning cost can be optimal—by inducing only the consumers with high-preference to actively learn the product features, the firm can price discriminate more effectively, and reduce the distortions on its choice of product features. We also discuss when it is optimal for the firm to provide a professional version of product by adding overly complicated features.

JEL Classification: D82, L11, M21

Key words: Second-Degree Price Discrimination, Screening, Consumer Learning.

1 Introduction

There are many products that require learning how to use. People sometimes need to understand product features and functionalities before purchasing products such as consumer electronics, personal computers, and software applications.[1] Consumers may not be able to use this kind of skill-based products without instructions. For example, although a consumer understands why Mathematica software is useful and places a high value on it, she may not make a purchase without knowing how to use it. Therefore, one central feature of this market is that the likelihood of purchase is affected by effort either by sellers or buyers.

When consumers have little or no experience with a product, a firm sometimes operates customer-help-centers to explain product features to consumers. Many firms also offer free trials of their products. Sometimes firms provide financial incentives for consumers to learn their product features. For example, university professors are frequently offered a learning opportunity of course management systems with monetary compensation. On the other hand, consumers expense their own resources to learn about the product. A consumer may spend a considerable time to know all the features of a product and learn how to use it.

A firm can adopt different approaches to sell skill-based products to consumers. One way is that a firm may expend resources to instruct consumers how to use its product. The other way is simply doing nothing and making consumers exert effort to learn the product features. For convenience, we will refer to the first case as 'firm-instructing' and the second case as 'consumer-learning'. In particular, in the consumer-learning case, the firm does not expend its resources instruct the consumers, but instead, it may have to indirectly compensate the consumers for their learning costs.

The objective of this paper is to see why and when one strategy dominates the other, thus providing a rationale for inducing consumer learning. We show that a firm can use a consumer's self-learning as a device to price-discriminate more effectively when a consumer's preference is private information. According to our result, when the learning cost is *small or large*, it is optimal for a firm to engage in the provision of instructions for use. When the learning cost is *intermediate*, by contrast, it is optimal for the firm to induce the consumer's self-learning. In addition, the optimal arrangement in the latter case is to induce only the high-preference consumers to engage in the learning activity. Since the firm's sales to a

[1] In addition, it is common that firms increase product features and functionalities greatly to enhance and differentiate their products. Brown and Carpenter (2000) and Thompson et al. (2005) point out that greater functionality can add too many features to products to the extent that consumers feel overwhelming and feature fatigue.

low-preference consumer is realized with a smaller probability under such an arrangement, the key trade-off in our paper is '*information extraction* vs. *loss of sales opportunities.*'

We employ a monopoly framework with second degree price discrimination. The firm can inform consumers of its product features and instructions for use, or induce the consumers to learn the product features.[2] A consumer's preference (type) can be either high or low in our model, and is her private information. As usual in the model of this type, a high-type consumer has an incentive to mimic a low-type consumer to reap information rent. When the firm provides instructions for the consumer, the optimal outcome for the firm is accompanied by the standard downward distortion for the low-type consumers.

Inducing the consumer to incur the learning cost has two merits. First, because a consumer knows her type, unlike when the firm provides instructions for the consumers, the consumers' self-learning incurs the learning cost only for the consumers of a particular type—as mentioned above, the firm makes only the high-type consumers to engage in active learning at the optimum. Second, and more importantly, by inducing only the high-type consumers to incur the learning cost, the firm can mitigate a consumer's the mimicry incentive, but only when the learning cost is intermediate.

When the firm induce only the high-type consumers to learn the product features, only the high-type consumers are compensated for their learning costs. The firm does so by offering an extra discount only for the high-type consumers (to be exact, only for the consumers choosing the product for the high-type). Such a discount discourages the high-type consumers from mimicking the low-type consumers. On the other hand, since the low-type consumers are not induced to make learning effort, the firm's sales to the low-type consumers are realized with a lower probability. Therefore, when the learning cost is small, the amount of discount for the high-type consumers is not large enough, and it is better for the firm to directly help the consumers for a higher sales probability.[3]

As the learning cost becomes larger, however, the incentive effect from the discount effect begins to dominate. When the learning cost is intermediate, while a high-type consumer can obtain information rent by misrepresenting her type, she is reluctant to do so because she then does not receive the discount. In other words, the consumer's self-learning allows the firm to create "countervailing incentives" in this range of the learning cost. As a result, the firm can price discriminate more effectively for rent extraction, which in turn allows it

[2] For simplicity, we assume that the learning costs are same whoever incurs the learning cost. Different learning costs do not change our results if the gap is not too large.

[3] In a different model, both the firms and the consumers may engage in learning at the optimum. It, however, will simply add more cases to our problem without altering the main insight.

to recover the distortion in the optimal choice of product features. We show that inducing self-learning can be optimal when the learning cost is in the intermediate range.

When the learning cost is large, the discount for the high-type consumers for their learning effort becomes so large that another incentive problem arises. In particular, if a low-type consumer purchases the product (without a learning effort), she will mimic the high-type to take advantage of the large discount. To prevent such a mimicry, the firm adds excessive features to the product for the high-type consumers. As the reverse incentive problem becomes an issue, it is optimal for the firm to provide instructions for use when the learning cost is large.

Our results mentioned above assume that, when inducing consumer self-learning, the firm offers a 'full menu' to a consumer regardless of the consumer's learning effort. We extend our analysis to the case in which the firm can offer limited menus by making the product for the high-type consumers overly complicated. That is to say, consumers cannot understand the product features for the high-type consumers without a learning effort. This is more like a professional or mania version of products. We show that when the learning cost is large, "the limited menu strategy" dominates "the full menu strategy".

The current paper is most closely related to the literature on second degree price discrimination in monopoly frameworks, pioneered by Mussa and Rosen (1978) and Maskin and Riley (1984). Studies that analyze situations in which consumers' product information include the following papers. Lewis and Sappington (1994) analyze the trade-off between price discrimination versus information rent. In their paper, the firm can provide more information about the product to a consumer for price discrimination, but such an activity makes the consumer realize her private information. Ottaviani and Prat (2001) show that committing to make a consumer's private information public can mitigate the consumer's misrepresenting incentive. Using a aggregate demand function, Johnson and Myatt (2006) extends the first paper to the case where the firm offers a product with different qualities.[4] Unlike our papers, the cost of information is incurred by the firm in these papers.[5] In our paper, the party that directly incurs the learning cost is endogenously determined.

Our paper is also related to the studies on "countervailing incentives." In their seminar work, Lewis and Sappington (1989) demonstrate that presence of countervailing incentives can improve the principal's welfare. Jullien (2000) provides a general analysis of type-dependent participation constraints with a continuum of types. The optimal mechanism with countervailing incentives and its benefit is applied in our paper. We show that induc-

[4] Bar-Issac et al. (2010) study informing consumers without price discrimination.

[5] Persico (2000) analyzes a situation in which the cost of information is incured by consumers.

ing consumers to incur learning costs, instead of the firm incurring such costs, generates countervailing incentives, which helps the firm extract a consumer's information rent.

The remainder of the paper is organized as follows. We set up the model in Section 2. In Section 3, the firm's optimal outcome from the instructed-learning case is presented. The firm's optimal outcome from the self-learning case is discussed in Section 4. In Section 5, we compare the two cases. In Section 6, we extend our analysis to the case where the firm offers limited menus by designing the high quality product overly complicated. We conclude in Section 7.

2 Model

A monopolist designs a product with new features $q \in \mathbb{R}_+$ and offers it at a price, $p \in \mathbb{R}_+$ to consumers. The population of consumers, for simplicity, is normalized to one. The consumer's preference toward the product is denoted by $i \in \{H, L\}$, where H (L) represents high (low) preference, and $\Delta \equiv H - L > 0$. With a probability φ_i, the consumer is type-i and $\sum_i \varphi_i = 1$. We assume that φ_i is not too small or too large that the firm does not exclude any type. The consumer's type is private information, but the probability distribution is common knowledge. As the revelation principle applies, q_i and p_i denote the quality and the lump sum price that the firm offers to type-i consumer.

The consumer values the product with a function $u(q_i, i)$ that is strictly increasing and concave in q_i with $u(0, i) = 0$. The value function also satisfies:

$$u^{\Delta}(q_i) \equiv u(q_i, H) - u(q_i, L) > 0 \quad \text{and} \quad u_q^{\Delta}(q_i) \equiv u_q(q_i, H) - u_q(q_i, L) > 0.$$

Adding q_i features to the product costs cq_i to the firm, where $c > 0$. The firm's profit and type-i's consumer's payoff from a transaction are respectively:

$$\Pi_i = p_i - cq_i \quad \text{and} \quad U_i = u(q_i, i) - p_i.$$

Consumers may not be informed of how to use the product features in the beginning. For simplicity, we assume that consumers do not make a purchase without knowing how to take advantage of the product features. Learning can increase the likelihood that a purchase occurs, and can be undertaken in the following two ways. The firm provides instructions for use ($\Psi = F$), or alternatively, induce the consumer to make a learning effort ($\Psi = C$). We denote by $e \in \{0, 1\}$ the instructing effort or learning effort level depending on $\Psi \in \{F, C\}$.

The probability of purchase is given by:

$$\gamma(e) = e + (1-e)\beta, \quad \beta \in (0,1).^{6}$$

A purchase can take place either when either instructing or learning effort is made ($e = 1$). When the consumer reads the instruction manual, she can understand product features with probability $\gamma = 1$. On the other hand, when the consumer learns how to use the product by accident ($e = 0$), the purchase can be made with probability $\gamma = \beta$. For example, a well-informed friend may explain the product features. The learning cost is given by se for all parties, where $s > 0$. We assume that the size of s is not too large that no learning is not optimal for the firm.

The timing is as follows:

(i) The consumer's type $i \in \{H, L\}$ is realized and privately observed by her.

(ii) The firm decides $\Psi \in \{F, C\}$.

(iii) The firm commits to the menu of its offers, $\{q_H,\, p_H : q_L,\, p_L\}$.

(iv) Depending on $\Psi \in \{F, C\}$, either the firm or the consumer makes the effort.

(v) A purchase takes place depending on $\gamma(e)$.

As a benchmark, we present the optimal outcome under full information. The first-best quality level, denoted by q_i^*, is characterized by the following equation:

$$u_q(q_i^*, i) = c, \quad i \in \{H, L\}.$$

Under full information, the optimal outcome satisfies "marginal benefit = marginal cost," with the perfect price discrimination, i.e., the firm leaves no consumer surplus: $p_i^* = u(q_i^*, i)$.

3 Firm-Instructing

In this section, we discuss the firm's optimal offers when the firm makes the instructing effort such as providing the product instructions for use or operating customer help centers. The problem is a standard non-linear pricing with screening and the firm solves the following problem:

$$\underset{q,p,e}{Max}\ \gamma(e)\left[\sum\nolimits_i \varphi_i\,(p_i - cq_i)\right] - se, \tag{$\mathcal{P}^F$}$$

subject to,

$$u(q_i, i) - p_i \geq 0, \quad i \in \{H, L\}, \tag{PC}$$

$$u(q_i, i) - p_i \geq u(q_j, i) - p_j, \quad i, j \in \{H, L\}. \tag{IC}$$

The first constraints, (PC), is the participation constraint for the consumer, and the second constraints, (IC), assures that the consumer's payoff is higher when she truthfully represents her type. The firm's decision of whether to make effort or not is independent of its choice of prices and quality. We restrict our attention to the relevant case where the effort always takes place ($e = 1$). That is, the learning cost s is not large enough.[7]

The following proposition presents the firm's optimal offer when it makes the effort for the consumer.

Proposition 1 *With $\Psi = F$, the the firm's optimal offers are characterized as follows:*

$$q_H^F = q_H^* \quad \textit{and} \quad q_L^F < q_L^*,$$

$$p_H^F < p_H^* \quad \textit{and} \quad p_L^F < p_L^*.$$

The result above is standard. The product features for type-H consumer is at the first best level, known as "efficiency at the top" in the literature, but the product features for type-L consumer is distorted downwards. As usual in the model of this type, type-H consumer has an incentive to mimic type-L consumer to reap information rent of $u^{\Delta}(q_L)$. The firm discourages such a mimicry by distorting the choice of product features for type-L consumer downwards. As a result, the firm must reduce both p_H and p_L from the first best levels, $p_H < p_H^*$ and $p_L < p_L^*$, resulting in an imperfect price discrimination.

4 Consumer-Learning

We now move on to the case in which the consumer makes a learning effort. The key difference from the case in the previous section comes from the fact that, when the consumer makes a learning effort, the firm can induce only a particular type to make a learning effort. In other words, the firm manipulates its offers such that the consumer of one type actively learns how to use the product by incurring the learning cost, whereas the other type just waits to learn how to use it by accident.

We first establish the following lemma.

[7] Formally, we assume $s \leq (1 - \beta) \left[\varphi_H \left(p_H^* - c q_H^* \right) + \varphi_L \left(p_L^F - c q_L^F \right) \right]$.

Lemma 1 *If $\Psi = C$ ever dominates $\Psi = F$, then the firm induces only type-H consumer to make a learning effort.*

It is not difficult to see that inducing both type-H and type-L consumer to incur the learning cost simply makes it more costly to the firm, compared to the case in which the firm makes the effort. Since type-L consumer gets zero consumer surplus from her purchase, she has no incentive to incur the learning cost to purchase the product. Therefore, to incentivize type-L consumer, the firm must provides her with a strictly positive consumer surplus. Such a surplus to type-L consumer, however, makes type-H consumer misrepresent her preference. As a result, the firm must provide the additional surplus to type-H consumer as well.

Similarly, it is suboptimal for the firm to induce only type-L consumer to make a learning effort. To induce type-L consumer's self-learning, the firm must compensate for the learning cost by decreasing the price, which attracts type-H consumer to misrepresent her type as type-L. Again, the firm will simply end up providing more consumer surplus under this arrangement than when it directly makes instructing effort.

Type-H consumer's problem when deciding whether to make a learning effort is:

$$\max\{u(q_H, H) - p_H - s, \quad \beta\,[u(q_H, H) - p_H]\},$$

implying that the firm's optimal offer for type-H consumer must satisfy:

$$u(q_H, H) - p_H \geq \frac{s}{1-\beta}. \tag{LC}$$

Since consumer surplus must be non-negative, the firm's maximization problem must satisfy the following participation constraints for the consumer:

$$u(q_H, H) - p_H - s \geq 0, \tag{PC_H}$$

$$\beta\,[u(q_L, L) - p_L] \geq 0. \tag{PC_L}$$

Finally, as the revelation principle applies in our model, the firm's offer must satisfy the incentive constraints for the consumer's truthful behavior:

$$u(q_H, H) - p_H - s \geq \max\left\{ \begin{array}{c} u(q_L, H) - p_L - s, \\ \beta\,[u(q_L, H) - p_L] \end{array} \right\}, \tag{IC_H}$$

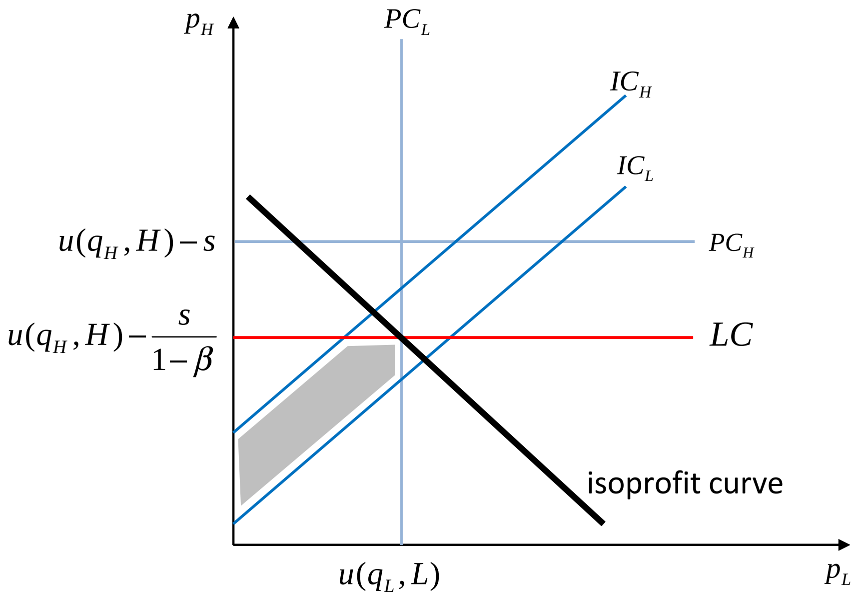


Figure 1: Countervailing Incentives - the Figure illustrate the case where (PC_L) and (LC) are only binding, and as a result, there is no distortion on both q_H and q_L.

$$\beta\left[u(q_L, L) - p_L\right] \geq \max \left\{ \begin{array}{c} u(q_H, L) - p_H - s, \\ \beta\left[u(q_H, L) - p_H\right] \end{array} \right\}. \tag{IC_L}$$

The constraints above assure that the consumer's payoff from truthfully representing her type (the left hand sides) is higher than her payoff from misrepresentation (the right hand sides). The right hand sides of (IC_H) and (IC_L) exhibit the consumer's choice of whether or not to learn the product features if she decides to misrepresent her type.

When making the consumer to incur the learning cost, the firm's problem is:

$$\underset{q,p}{Max}\, \varphi_H\left(p_H - cq_H\right) + \beta\varphi_L\left(p_L - cq_L\right), \tag{$\mathcal{P}^C$}$$

subject to $(LC), (PC_H), (PC_L), (IC_H)$, and (IC_L). Notice that (LC) implies (PC_H).

There are five different regimes for the optimal outcome when the consumer makes the learning effort. The analysis can be illustrated with the help of Figure 1. Given any q_H and q_L, we draw each constraint in p_H, p_L-space. Note that the isoprofit curve is linear and downward sloping with the slope equal to $\frac{\beta\varphi_L}{\varphi_H}$. First, when s is zero or small enough, as in a standard adverse selection model, it can be easily seen that only (PC_L) and (IC_H) are binding. Second, as s becomes larger, (LC) meets (PC_L) and (IC_H) simultaneously, and three constraints together are binding. Third, for greater s, only (LC) and (PC_L) bind. Fourth, (LC) meets (PC_L) and (IC_L) simultaneously. Last, as s becomes large enough, only (LC) and (IC_L) bind. Below, however, we present the results as if there are three different regimes for simplicity. This way also highlights the intermediate case where none of incentive constraints binds.

To characterize the optimal outcomes in each regime, we first present the following cutoff levels of the learning cost.

Definition 1 *Let* $\underline{s} \equiv (1-\beta)u^{\Delta}\left(q_L^*\right)$ *and* $\overline{s} \equiv (1-\beta)u^{\Delta}\left(q_H^*\right)$.

The following proposition characterizes the firm's optimal offers.

Proposition 2 *With* $\Psi = C$, *the firm's optimal offers are characterized as follows:*

- *When* $s < \underline{s}$: $q_H^C = q_H^F$ *and* $q_L^C \lesseqgtr q_L^F$, $p_H^C \gtreqless p_H^F$ *and* $p_L^C \lesseqgtr p_L^F$.
- *When* $s \in [\underline{s}, \overline{s}]$: $q_i^C = q_i^* \ \forall i$, $p_H^C = p_H^* - \frac{s}{1-\beta}$, *and* $p_L^C = p_L^*$.
- *When* $s > \overline{s}$: $q_H^C > q_H^F$ *and* $q_L^C = q_L^*$, $p_H^C < p_H^F$ *and* $p_L^C \gtreqless p_L^F$.

When the learning cost is sufficiently small, type-L consumer's lack of learning effort makes the distortion in the provision of product features to her larger than the one with $\Psi = F$ $(q_L^C < q_L^F)$. Yet, type-H consumer's information rent $u^{\Delta}(q_L)$ is large enough that the firm does not need to compensate her for the learning cost, and consequently, $p_H^C > p_H^F$. Since the product features for type-L consumer is lower compared to when $\Psi = F$, the price is lower $(p_L^C < p_L^F)$. As the learning cost increases, the firm must incentivize type-H consumer to make a learning effort by providing a discount $(p_H^C < p_H^F)$. This extra consumer surplus is provided only when type-H consumer truthfully represent her type, allowing the firm to recover some of distortion in the optimal choice of product features for type-L consumer $(q_L^C > q_L^F)$. As a result, the price for type-L consumer becomes higher than when the firm makes instructing effort $(p_L^C > p_L^F)$.

As the learning cost increases, while it becomes more costly to induce type-H consumer to make a learning effort, her private information becomes less of a problem. Within an intermediate range, the firm's extra discount to type-H consumer to induce her learning effort is large enough that the consumer no longer has an incentive to misrepresent her type as type-L. That is, the learning cost is large enough that type-L consumer also has an incentive to misrepresent her type as type-H, and the misrepresenting incentives of opposite directions cancel out each other, i.e., "countervailing incentives" arise. Consequently, the firm does not need to distort the quality choice in its optimal offer to extract the consumer's surplus associated with her private information ($q_i^C = q_i^*$, $\forall i$). Within this range of the learning cost, although the firm must give type-H consumer a discount for her learning effort, the prices are not distorted due to the consumer's private information.

As the learning cost becomes yet larger, so becomes the discount to type-H consumer, which leads to a "reverse incentive" problem. In this regime, type-H consumer has no misrepresenting incentive, but type-L consumer has such an incentive when she finds the product by chance. As a result, the optimal quality choice for type-H consumer is distorted upward ($q_H^C > q_H^F = q_H^*$). If type-L misrepresents her type as type-H, then her purchasing product must have excessive product features, which in turn discourages her misrepresentation. This effect contributes positively to the price for type-L consumer. At the same time, the firm must also give type-L consumer a discount to prevent her mimicry. As a result, the price for type-L consumer can be higher or lower compared to the case with $\Psi = F$ ($p_L^C \gtreqless p_L^*$).

In summary, making the consumer learning the product features, instead of the firm making instructing effort, brings about different incentive problems according to the size of the learning cost. In the next section, we examine pros and cons of each sales strategy, and endogenize the firm's choice of $\Psi \in \{F, C\}$ to determine its optimal sales strategy.

5 Optimal Sales Strategy

By directly instructing the consumer ($\Psi = F$), the firm can sell its product with a higher probability, since a learning always takes place. This, however, implies that the learning cost is always incurred under such an arrangement, regardless of the consumer's type. Moreover, when the firm makes the instructing effort, it must always provide type-H consumer with strictly positive information rent.

The consumer's self-learning ($\Psi = C$) brings more flexibility to the firm — it allows the firm to incur the learning cost depending on the consumer's type. Such a flexibility allows the firm to partially save the learning cost, but more importantly, it has a strategic benefit. In particular, when the learning cost is not too small or large, the consumer's self-learning brings about the countervailing incentives, which enables the firm to extract the consumer's information rent.

It is shown in the appendix that the firm's expected profit with $\Psi = F$ is linearly decreasing in s, whereas the firm's expected profit with $\Psi = C$ is non-increasing and weakly concave in s. As a result, the firm's choice of $\Psi \in \{F, C\}$ is non-monotonic in s. In the following proposition, we present the firm's optimal sales strategy depending on the learning cost.

Proposition 3 *Suppose $\Delta \equiv H - L$ is large enough:*

- *When s is small, the firm chooses $\Psi = F$.*
- *When s is intermediate, the firm chooses $\Psi = C$.*
- *When s is large, the firm chooses $\Psi = F$.*

When the learning cost is small, instructing the consumer is more attractive to the firm for a higher probability that a purchase takes place. When the learning cost is in the intermediate range, the consumer's self-learning enables the firm to extract the consumer's information rent, leading to more efficient price discrimination. This effect dominates when the difference in the consumer's preferences is large enough, and consequently, the firm prefers the consumer's self-learning. When the learning cost is large, however, such a strategic merit vanishes since the consumer's self-learning brings about the reverse incentive problem. In such a case, the same discount to type-H consumer for her learning effort must also be provided to type-L consumer. Consequently, it becomes too costly to induce the consumer to make a learning effort, and the firm prefers instructing the consumer to making the consumer's self-learning.

Corollary 1 *The way that the firm's offer deviates from the social optimum is different depending on $\Psi \in \{F, C\}$. With $\Psi = F$, $q_L^F < q_L^*$. With $\Psi = C$, even if $q_H^C = q_H^*$ and $q_L^C = q_L^*$, there is the loss of sales to type-L consumer. Also, $q_H^C > q_H^*$ is also possible .*

Let us now briefly discuss the implications for social welfare. When $\Psi = F$, the firm's choice of price and product features is the standard outcome in the monopolistic screening model. The deviation from the social optimum is the downward distortion on the number of product features for type-L consumer. The product for type-L consumer has a smaller capability than the socially optimal level. On the other hand, when $\Psi = C$, we may not have any distortion on the number of product features. In such a case, nevertheless, there is still the loss of efficiency, which is that the type-L consumer's purchase does not occur. In addition, with $\Psi = C$, the firm may offer the product with too many features type-H consumer. In this case, we have not only the loss of sales for type-H consumer, but the upward distortion on the number of product features.

6 Limited Menu Strategy

So far, we have considered the case in which the consumer is able to learn the product features either by making a learning effort or by accident. However, firms often offer a professional version of products that make it very hard for consumers to learn the instructions without a significant effort. Similarly, a firm may add considerably complicated features to the product. Consumers are not able to understand these features at all without a learning effort.

To fix the idea, let us distinguish two different strategies. The first one is what has been discussed—offering $\{q_H, p_H : q_L, p_L\}$ regardless of the consumer's learning effort. We refer to this as "the full menu strategy." The second one is the strategy that makes $\{q_H, p_H\}$ available only for the consumer who makes a learning effort ($e = 1$). We refer to this as "the limited menu strategy ($\Psi = \widetilde{C}$).

Analogously speaking, the firm can set up different purchasing channels for the consumer: "The learning channel" through which $\{q_H, p_H : q_L, p_L\}$ is offered and the consumer must make a learning effort to get to this channel, and "the regular channel" through which only $\{q_L, p_L\}$ is offered. Figure 2 below illustrates the difference between two cases: $\Psi = C$ and $\Psi = \widetilde{C}$.[8]

As illustrated in Figure 1, the limited menu strategy has two main differences from the full menu strategy. First, type-H consumer, if she decides not to purchase through the learning channel, must take the offer for type-L consumer. Therefore, the learning constraint becomes:

$$u(q_H, H) - p_H - s \geq \beta \left[u(q_L, H) - p_L\right]. \qquad (\widetilde{LC})$$

Second, type-L consumer must incur the learning cost if she decides to mimic type-H. Therefore, instead of (IC_L) in the full menu strategy, the firm must satisfy the following incentive constraint for type-L consumer's truthful behavior:

$$\beta \left[u(q_L, L) - p_L\right] \geq u(q_H, L) - p_H - s. \qquad (\widetilde{IC}_L)$$

The firm maximizes its expected payoff in $(\mathcal{P}^C)$, subject to (PC_H), (PC_L), $(\widetilde{LC})$, (IC_H) and $(\widetilde{IC}_L)$. To characterize the optimal outcome in each regime, we first present the

[8] One can consider another case that the firm may offer $\{q_H, p_H\}$ through the learning channel, and offer $\{q_L, p_L\}$ through the regular channel. In this case, the only difference is that (IC_H) does not exist in the principal's problem below. Our main point, Proposition5, still remains. We do not analyze this case because this case is not realistic in that type-H consumers cannot mimic type-L consumers even with the learning effort.

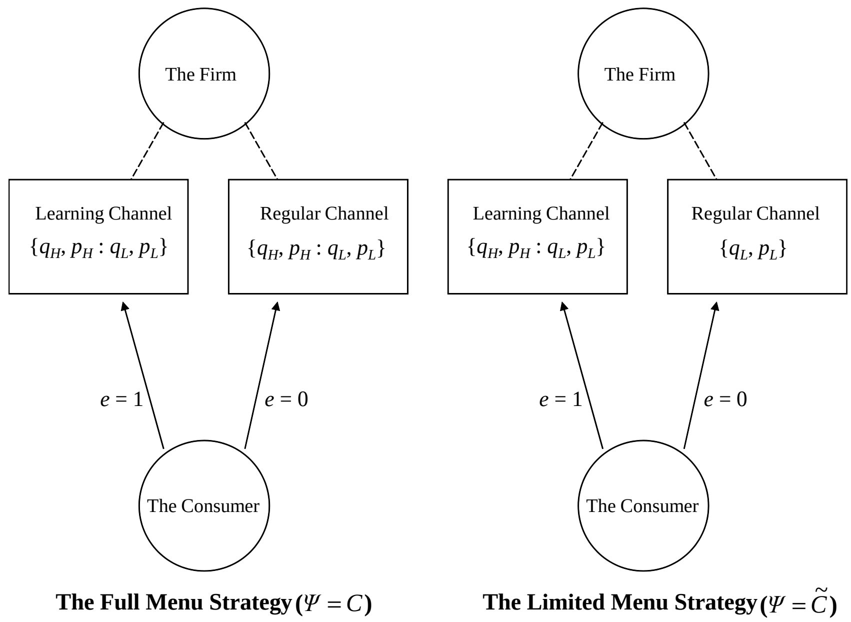


Figure 2: Inducing Consumer Search with Different Strategies

following cutoff levels of the learning cost.[9]

Definition 2 *Let* $\underline{s}^{\widetilde{c}} \equiv u_q^{\Delta}(q_L^*) - \beta u^{\Delta}(q_L^*)$.

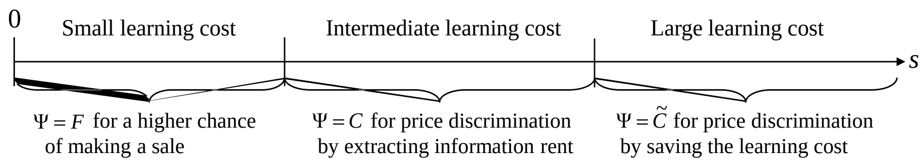
The firm's optimal offers are characterized in the next proposition.

Proposition 4 *With* $\Psi = \widetilde{C}$, *the firm's optimal offers with the limited menu strategy are characterized as follows:*

- *When* $s < \underline{s}^{\widetilde{c}}$*:* $q_H^{\widetilde{C}} = q_H^F$ *and* $q_L^{\widetilde{C}} \lesseqgtr q_L^F$.
- *When* $s \geq \underline{s}^{\widetilde{c}}$*:* $q_i^C = q_i^* \ \forall i$.

The limited menu strategy has a trade-off with the full menu strategy. The incentive constraint for type-L consumer becomes loose because it is harder for type-L consumer to mimic type-H consumer. Indeed, $(\widetilde{IC}_L)$ is never binding as will be shown in the appendix. This implies, differently from the full menu strategy, that the limited menu strategy does not have the upward incentive distortion problem when a learning cost is large enough.

[9] We have three different regimes for the optimal outcome. However, we present the results as if there are two different regime for simplicity and for consistency following Proposition 2.



On the other hand, the learning constraint can be more demanding with the limited menu strategy. This is because, under the limited menu strategy, type-H consumer may have less incentives to make the learning effort when the product for type-L consumer is still attractive.[10] Consequently, the firm may provide a larger discount to type-H consumer with the limited menu strategy. This trade-off leads to a change in the firm's optimal strategy in the following way.

Proposition 5 *Suppose $\Delta \equiv H - L$ is large enough and c is small enough:*

- *When s is small, the firm chooses $\Psi = F$.*
- *When s is intermediate, the firm chooses $\Psi = C$.*
- *When s is large, the firm chooses $\Psi = \widetilde{C}$.*

As presented in the proposition above, when the learning cost is large, the limited menu strategy becomes optimal for the firm. Again, the full menu strategy leads to type-L consumer's incentive to mimic type-H for a large learning cost. In order to keep type-L consumer from such a mimicry, the firm must provide a discount for type-H consumer's learning effort regardless of the consumer's type when the learnzing cost is large. The limited menu strategy limits flexibility in the firm's pricing strategies for different levels of the learning cost, but such a limit, in turn, restricts the consumer's misrepresenting incentives. For large learning costs, therefore, the limited menu strategy allows the firm to avoid giving a discount to type-L consumer without a distortion. Finally, the limited menu strategy dominates the case where the firm makes the instructing effort because the firm can save the costly effort for type-L consumer. The result is illustrated in Figure 3 below.

[10] In other words, q_L^* is large enough as shown in the appendix. This is the case when c is small enough. On the other hand, when c is not small, $\Psi = \widetilde{C}$ can dominate $\Psi = C$.

7 Conclusion

In this paper, we have shown that a firm can employ consumer's self-learning as a device for second degree price discrimination. When a consumer's preference is private information, the firm can extract the consumer's surplus more efficiently by inducing the consumer to make a learning effort. According to our result, when the learning cost is small or large, it is optimal for the firm to instruct the consumer, by directly bearing the learning cost. For intermediate learning costs, the firm prefers inducing the consumer with a higher preference to bear the learning cost and compensate it through the price. We have also shown that offering a limited menu, equivalently a professional version, can dominate even when the learning cost is large.

For expositional purpose, we made two simplifying assumptions. First, we assumed that the consumer is risk-neutral. If the consumer is risk averse, then the firm must compensate type-L consumer for the risk she takes by not learning. Although our qualitative result will still hold, the range of the learning cost in which inducing consumer's self-learning is optimal will become narrower. Second, our result can be extended to the case with a continuum of the consumer's types. Then, unlike the two-type case, inducing consumer learning allows the firm to fully extract the consumer's information rent only at a particular level of learning cost. Nevertheless, the firm will still prefer consumer learning in a intermediate range of the learning cost. Also, with a continuum of types, the distortion on product features becomes more sensitive to the curvature of the learning cost.[11]

[11] See Maggi and Rodriguez-Clare (1995) for a formal analysis related to this issue.

Appendix: Proofs

Proof of Proposition 1.

Since s does not play any role, only (IC) for type-H consumer and (PC) for type-L consumer are binding as in a standard screening problem:

$$p_H = u(q_H, H) - u^{\Delta}(q_L) \quad \text{and} \quad p_L = u(q_L, L). \tag{$A1$}$$

Substituting for p_H and p_L in the objective function in $(\mathcal{P}^F)$, and we solve:

$$\underset{q_H, q_L}{Max}\, \varphi_H \left[u(q_H, H) - u^{\Delta}(q_L) - cq_H\right] + \varphi_L \left[u(q_L, L) - cq_L\right] - s. \tag{$A2$}$$

The first order conditions give:

$$u_q(q_H^F, H) = c \quad \text{and} \quad u_q(q_L^F, L) = c + \frac{\varphi_H}{\varphi_L} u_q^{\Delta}(q_L^F),$$

implying that $q_H^F = q_H^*$ and $q_L^F < q_L^*$. From $(A1)$, $p_H^F < p_H^*$ and $p_L^F < p_L^*$.$\blacksquare$

Proof of Lemma 1.

We show that, with $\Psi = C$, (i) the case in which the firm induces a learning effort from the consumer regardless of her type and (ii) the case in which the firm only induces type-L consumer's learning effort are both dominated by the firm's optimal outcome with $\Psi = F$. First, suppose, with $\Psi = C$, the firm induces a learning effort from both types. Then, the following learning constraint must be satisfied:

$$u(q_i, i) - p_i - s \geq \beta \left[u(q_i, i) - p_i\right], \quad i \in \{H, L\},$$

which can be rewritten as:

$$u(q_i, i) - p_i \geq \frac{s}{1-\beta}, \quad i \in \{H, L\}. \tag{$A3$}$$

The consumer's participation constraint $u(q_i, i) - p_i - s \geq 0$ is implied by $(A3)$ regardless of her type. The constraints that induce the consumer's truthful representation of her type are:

$$u(q_i, i) - p_i - s \geq \max \left\{ \begin{array}{c} u(q_j, i) - p_j - s, \\ \beta \left[u(q_j, i) - p_j\right] \end{array} \right\}, \quad i, j \in \{H, L\}. \tag{$A4$}$$

The RHS of $(A4)$ exhibits the consumer's choice of whether or not to exert a learning effort if she decides to misrepresents her type (off the equilibrium path). To simplify $(A4)$, we first present the following lemma.

Lemma 2 *Suppose* $(A3)$ *holds. Then, the inequality below must hold if* $(A4)$ *for type-*i *consumer is binding:*

$$u(q_j, i) - p_j \geq \frac{s}{1-\beta}, \quad i, j \in \{H, L\}.$$

Proof. Suppose $u(q_j, i) - p_j < \frac{s}{1-\beta}$, implying that $u(q_j, i) - p_j - s < \beta [u(q_j, i) - p_j]$ in the RHS of $(A4)$. Then, with binding $(A4)$ for type-i consumer can be rewritten as:

$$u(q_i, i) - p_i - \frac{s}{1-\beta} = \beta \left[u(q_j, i) - p_j - \frac{s}{1-\beta} \right],$$

which is a contradiction since the LHS of the equation is positive by $(A3)$, but the RHS is negative. ■

Lemma 2 implies that RHS of $(A4)$ is $u(q_j, i) - p_j - s$. Therefore, $(A4)$ becomes the same as (IC) in the case with $\Psi = F$, since s cancels out with each other in both sides of $(A4)$. The firm's problem then is written as:

$$\underset{q_i, p_i}{Max} \sum\nolimits_i \varphi_i \left(p_i - cq_i \right),$$

subject to

$$u(q_i, i) - p_i \geq \frac{s}{1-\beta}, \quad i \in \{H, L\}, \tag{A5}$$

$$u(q_i, i) - p_i \geq u(q_j, i) - p_j, \quad i, j \in \{H, L\}. \tag{A6}$$

Compared to the case with $\Psi = F$, there are two differences. First, the learning cost s is transferred to the consumer, and second, the consumer's reservation payoff is $\frac{s}{1-\beta}$. As usual, $(A5)$ for type-L and $(A6)$ for type-H consumer are binding at the optimum, and we have expression for p_i, $i \in \{H, L\}$, from these binding constraints. After substituting for the prices in the the firm's objective function, the optimization problem becomes:

$$\underset{q_H, q_L}{Max} \quad \varphi_H \left[u(q_H, H) - u^{\Delta}(q_L) - cq_H \right] + \varphi_L \left[u(q_L, L) - cq_L \right] - \frac{s}{1-\beta}. \tag{A7}$$

Directly comparing $(A7)$ to $(A2)$ shows that the firm's profit with $\Psi = C$ is strictly lower than its profit with $\Psi = F$.

Next, suppose, with $\Psi = C$, the firm induces a learning effort only from type-L consumer. Then, the following learning constraint must be satisfied:

$$u(q_L, L) - p_L \geq \frac{s}{1-\beta}. \tag{A8}$$

The constraints that induce the consumer's truthful representation of her type are:

$$\beta\left[u(q_H, H) - p_H\right] \geq \max\left\{\begin{array}{c} u(q_L, H) - p_L - s, \\ \beta\left[u(q_L, H) - p_L\right] \end{array}\right\}, \tag{A9}$$

$$u(q_L, L) - p_L - s \geq \max\left\{\begin{array}{c} u(q_H, L) - p_L - s, \\ \beta\left[u(q_H, L) - p_H\right] \end{array}\right\}. \tag{A10}$$

Lemma 3 *Suppose $u(q_H, H) - p_H < \frac{s}{1-\beta}$. Then, the inequality below must hold:*

$$u(q_L, H) - p_L < \frac{s}{1-\beta}.$$

Proof. Suppose $u(q_L, H) - p_L \geq \frac{s}{1-\beta}$, implying that $u(q_L, H) - p_L - s \geq \beta\left[u(q_L, H) - p_L\right]$ in the RHS of ($A9$). Then, ($A9$) with a simple manipulation gives:

$$\beta\left[u(q_H, H) - p_H - \frac{s}{1-\beta}\right] \geq u(q_L, H) - p_L - \frac{s}{1-\beta},$$

which is a contradiction since the LHS is negative, but the RHS is positive. ■

Lemma 3 implies that the RHS of ($A9$) is $\beta\left[u(q_L, H) - p_L\right]$. Also by Lemma 2, RHS of ($A10$) is $u(q_H, L) - p_L - s$. Therefore, ($A9$) and ($A10$) become the same as (IC) in the case with $\Psi = F$. For type-L consumer, ($A8$) implies that the constraint for her participation, $u(q_L, L) - p_L - s \geq 0$, is automatically satisfied. By Lemma 3, ($A9$) is written as $u(q_H, H) - p_H \geq u(q_L, H) - p_L$, which implies that participation constraint for type-H consumer $u(q_H, H) - p_H \geq 0$ is automatically satisfied. Therefore, the firm's problem is written as:

$$\underset{q,p}{Max}\ \beta\varphi_H\left(p_H - cq_H\right) + \varphi_L\left(p_L - cq_L\right),$$

subject to

$$u(q_L, L) - p_L \geq \frac{s}{1-\beta}, \tag{A11}$$

$$u(q_i, i) - p_i \geq u(q_j, i) - p_j, \quad i, j \in \{H, L\}. \tag{A12}$$

Again, by Lemma 2 and 3, $(A9)$ and $(A10)$ become the constraints in $(A12)$. It can be easily shown that $(A11)$ and $(A12)$ for type-H are binding, and $(A12)$ for type-L consumer is slack. By substituting for p_L and p_H in the objective function, the firm's problem becomes:

$$\underset{q_H, q_L}{Max} \quad \beta\varphi_H \left[u(q_H, H) - u^{\Delta}(q_L) - cq_H - \frac{s}{1-\beta}\right] + \varphi_L \left[u(q_L, L) - cq_L - \frac{s}{1-\beta}\right]. \quad (A13)$$

Clearly, the firm's profit in $(A13)$ is even smaller than it's profit in $(A7)$. ■

Proof of Proposition 2.

First, the following two lemmas establish that (IC_H) and (IC_L) become:

$$u(q_H, H) - p_H \geq u(q_L, H) - p_L \text{ and } u(q_L, L) - p_L \geq u(q_H, L) - p_H.$$

Lemma 4 *Suppose (LC) holds. Then, the inequality below must hold if (IC_H) is binding:*

$$u(q_L, H) - p_L \geq \frac{s}{1-\beta}.$$

Proof. Suppose $u(q_L, H) - p_L < \frac{s}{1-\beta}$, implying that $u(q_L, H) - p_L - s < \beta\left[u(q_L, H) - p_L\right]$ in the RHS of (IC_H) Then, binding (IC_H) can be rewritten as:

$$u(q_H, H) - p_H - \frac{s}{1-\beta} = \beta\left[u(q_L, H) - p_L - \frac{s}{1-\beta}\right],$$

which is a contradiction since the LHS is positive by (LC), but the RHS is negative. ■

The lemma above implies that RHS of (IC_H) is $u(q_L, H) - p_L - s$. Therefore, (IC_H) is rewritten as $u(q_H, H) - p_H \geq u(q_L, H) - p_L$.

Lemma 5 *Suppose $u(q_L, L) - p_L < \frac{s}{1-\beta}$. Then, the inequality below must hold:*

$$u(q_H, L) - p_H < \frac{s}{1-\beta}.$$

Proof. Suppose $u(q_H, L) - p_H \geq \frac{s}{1-\beta}$, implying that $u(q_H, L) - p_H - s \geq \beta\left[u(q_H, L) - p_H\right]$ in the RHS of (IC_L). Then, (IC_L) with a simple manipulation gives:

$$\beta\left[u(q_L, L) - p_L - \frac{s}{1-\beta}\right] \geq u(q_H, L) - p_H - \frac{s}{1-\beta},$$

which is a contradiction since the LHS is negative, but the RHS is positive. ■

The lemma above implies that RHS of (IC_L) is $\beta\left[u(q_L,H)-p_L\right]$. Therefore, (IC_L) is rewritten as $u(q_L,L)-p_L\geq u(q_H,L)-p_H$.

Since (LC) implies (PC_H), the Lagrangian of the firm's problem can be written as:

$$\begin{aligned}
\mathcal{L} \;=\;& \varphi_H\left(p_H-cq_H\right)+\beta\varphi_L\left(p_L-cq_L\right)\\
&+\lambda_1\left[u(q_H,H)-p_H-\frac{s}{1-\beta}\right]\\
&+\lambda_2[u(q_L,L)-p_L]\\
&+\lambda_3\left[u(q_H,H)-p_H-u(q_L,H)+p_L\right]\\
&+\lambda_4\left[u(q_L,L)-p_L-u(q_H,L)+p_H\right].
\end{aligned}$$

The first order conditions are:

$$\frac{\partial\mathcal{L}}{\partial p_H}=\varphi_H-\lambda_1-\lambda_3+\lambda_4=0, \tag{A14}$$

$$\frac{\partial\mathcal{L}}{\partial p_L}=\beta\varphi_L-\lambda_2+\lambda_3-\lambda_4=0, \tag{A15}$$

$$\frac{\partial\mathcal{L}}{\partial q_H}=-\varphi_Hc+(\lambda_1+\lambda_3)u_q(q_H,H)-\lambda_4u_q(q_H,L)=0, \tag{A16}$$

$$\frac{\partial\mathcal{L}}{\partial q_L}=-\beta\varphi_Lc+(\lambda_2+\lambda_4)u_q(q_L,L)-\lambda_3u_q(q_L,H)=0, \tag{A17}$$

$$\frac{\partial\mathcal{L}}{\partial\lambda_1}=u(q_H,H)-p_H-\frac{s}{1-\beta}\geq 0,\quad \lambda_1\frac{\partial\mathcal{L}}{\partial\lambda_1}=0, \tag{A18}$$

$$\frac{\partial\mathcal{L}}{\partial\lambda_2}=u(q_L,L)-p_L\geq 0,\quad \lambda_2\frac{\partial\mathcal{L}}{\partial\lambda_2}=0, \tag{A19}$$

$$\frac{\partial\mathcal{L}}{\partial\lambda_3}=u(q_H,H)-p_H-u(q_L,H)+p_L\geq 0,\;\; \lambda_3\frac{\partial\mathcal{L}}{\partial\lambda_3}=0, \tag{A20}$$

$$\frac{\partial\mathcal{L}}{\partial\lambda_4}=u(q_L,L)-p_L-u(q_H,L)+p_H\geq 0,\;\; \lambda_4\frac{\partial\mathcal{L}}{\partial\lambda_4}=0. \tag{A21}$$

As will be shown below, the first regime, $s<\underline{s}$ is divided into two sub-regimes, $s<\underline{\underline{s}}$ and $s\in[\underline{\underline{s}},\underline{s})$, and $s>\overline{s}$ is also divided into two sub-regimes, $s\in(\overline{s},\overline{\overline{s}}]$ and $s>\overline{\overline{s}}$. The following two lemmas establish the binding constraints in each case.

Lemma 6 *$\lambda_3\lambda_4=0$, i.e., (IC_H) and (IC_L) cannot be simultaneously binding.*

Proof. Suppose $\lambda_3 > 0$ and $\lambda_4 > 0$. Then, from $(A14)$ and $(A15)$, $\lambda_3 = \varphi_H - \lambda_1 + \lambda_4$ and $\lambda_4 = \beta\varphi_L - \lambda_2 + \lambda_3$. Also, from $(A16)$ and $(A17)$, $\lambda_3 = \frac{\varphi_H c}{u_q(q_H,H)} - \lambda_1 + \frac{u_q(q_H,L)}{u_q(q_H,H)}\lambda_4$ and $\lambda_4 = \frac{\beta\varphi_L c}{u_q(q_L,L)} - \lambda_2 + \frac{u_q(q_L,H)}{u_q(q_L,L)}\lambda_3$. These equations imply that $u_q(q_H, H) = u_q(q_L, L) = u_q(q_H, L) = u_q(q_L, H)$, which is not possible. $\blacksquare$

Lemma 7 *If $\lambda_4 = 0$ ($\lambda_3 = 0$), then $\lambda_2 > 0$ ($\lambda_1 > 0$) and $q_H^C = q_H^*$ ($q_L^C = q_L^*$). In other words, (i) if (IC_H) is non-binding, then (LC) is binding and $q_H^C = q_H^*$, and (ii) if (IC_L) is non-binding, then (PC_H) is binding and $q_L^C = q_L^*$.*

Proof. Suppose $\lambda_4 = \lambda_2 = 0$. $(A17)$ gives $\lambda_3 < 0$, which is a contradiction. From $(A14)$ and $(A16)$ with $\lambda_4 = 0$, we obtain $u_q(q_H, H) = c$, and thus $q_H^C = q_H^*$. Similarly, $\lambda_3 = \lambda_1 = 0$ yields $\lambda_4 < 0$ in $(A16)$, which is a contradiction. Also, $(A15)$ and $(A17)$ with $\lambda_3 = 0$, we obtain $u_q(q_L, L) = c$, and thus $q_L^C = q_L^*$. $\blacksquare$

By Lemma 6 and 7, we can confine our attention to the five cases below. Case I, in which $\lambda_1 = 0$, $\lambda_2 > 0$, $\lambda_3 > 0$ and $\lambda_4 = 0$ ((PC_L) and (IC_H) are binding), Case II, in which $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 > 0$ and $\lambda_4 = 0$ ((LC), (PC_L) and (IC_H) are binding), Case III, in which $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 = 0$ and $\lambda_4 = 0$ ((LC) and (PC_L) are binding), Case IV, in which $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 = 0$ and $\lambda_4 > 0$ ((LC), (PC_L) and (IC_L) are binding), and Case V, in which $\lambda_1 > 0$, $\lambda_2 = 0$, $\lambda_3 = 0$ and $\lambda_4 > 0$ ((LC) and (IC_L) are binding). We show that Case I and II belong to the regime of $s < \underline{s}$, Case III to the regime of $s \in [\underline{\underline{s}}, \underline{s})$, and Case IV and V to the regime of $s > \overline{s}$.

$$\begin{aligned}
\underline{\underline{q_L}} \text{ by } u_q(\underline{\underline{q_L}}, L) &= c + \frac{\varphi_H}{\beta\varphi_L} u_q^{\Delta}(\underline{\underline{q_L}}), && (A22)\\
\underline{q_L} \text{ by } u^{\Delta}(\underline{q_L}) &= \frac{s}{1-\beta}, && (A23)\\
\overline{q_H} \text{ by } u^{\Delta}(\overline{q_H}) &= \frac{s}{1-\beta}, \text{ and} && (A24)\\
\overline{\overline{q_H}} \text{ by } u_q(\overline{\overline{q_H}}, H) &= c - \frac{\beta\varphi_L}{\varphi_H} u_q^{\Delta}(\overline{\overline{q_H}}). && (A25)
\end{aligned}$$

Case I: $\lambda_1 = 0$, $\lambda_2 > 0$, $\lambda_3 > 0$ and $\lambda_4 = 0$. Suppose s is close to zero. Then, the problem becomes a standard screening problem in which only (IC_H) and (PC_L) are binding. From $(A14)$ and $(A16)$ with $\lambda_4 = 0$, we have $u_q(q_H, H) = c$, and thus $q_H^C = q_H^F$ $(= q_H^*)$. To find q_L^C, we insert $\lambda_2 = \varphi_H + \beta\varphi_L$ into $(A17)$. We have:

$$u_q(q_L, L) = c + \frac{\varphi_H}{\beta\varphi_L} u_q^{\Delta}(q_L), \tag{4}$$

which implies $q_L^C = \underline{\underline{q_L}} < q_L^F$. Accordingly, $p_H^C = u(q_H^*, H) - u_q^{\Delta}(\underline{\underline{q_L}}) > p_H^F$ and $p_L^C = u(\underline{\underline{q_L}}, L) < p_L^F$. The non-binding constraint associated with λ_1 is written as:

$$u^{\Delta}(\underline{\underline{q_L}}) - \frac{s}{1-\beta} > 0.$$

As s increases, the constraint will eventually be binding at $s = \underline{\underline{s}}$, where $\underline{\underline{s}} = (1-\beta)u^{\Delta}(\underline{\underline{q_L}})$.

Case II: $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 > 0$ and $\lambda_4 = 0$. As s becomes $\underline{\underline{s}}$, the constraint linked to λ_1, (LC), begins to bind as well. In this case, the binding $(A18)$, $(A19)$, and $(A20)$ give:

$$u^{\Delta}(q_L) = \frac{s}{1-\beta},$$

which implies $q_L^C = \underline{q_L} \lesseqgtr q_L^F$, depending on the size of s. From $(A14)$ and $(A16)$ with $\lambda_4 = 0$, we have $q_H^C = q_H^F$ $(= q_H^*)$. From the binding constraints, we have $p_H^C = u(q_H^*, H) - \frac{s}{1-\beta} \gtreqless p_H^F$ and $p_L^C = u(\underline{q_L}, L) \lesseqgtr p_L^F$. Case II is now valid when $s \in [\underline{\underline{s}}, \underline{s})$, where $\underline{s} = (1-\beta)u^{\Delta}(q_L^*)$. If s becomes greater than $\underline{s}$, the constraint linked to λ_3, (IC_H), becomes no longer binding.

Case III: $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 = 0$ and $\lambda_4 = 0$. Only (LC) and (PC_L) are binding in this sub-regime. As in the above two cases, $\lambda_4 = 0$ implies that $q_H^C = q_H^*$. From $(A15)$ and $(A17)$ with $\lambda_3 = 0$ we have $q_L^C = q_L^*$. From the binding constraints, we have $p_H^C = u(q_H^*, H) - \frac{s}{1-\beta} \left(= p_H^* - \frac{s}{1-\beta}\right)$ and $p_L^C = u(q_L^*, L) = p_L^*$. The non-binding constraint related with λ_4, (IC_L), is written as:

$$u^{\Delta}(q_H^*) > \frac{s}{1-\beta}.$$

As s becomes larger, this constraint will bind at $s = \overline{s}$, where $\overline{s} = (1-\beta)u^{\Delta}(q_H^*)$.

Case IV: $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 = 0$ and $\lambda_4 > 0$. As $s > \overline{s}$, we consider the case where the constraint linked to λ_4, (IC_L), begins to bind. Note that q_H^C is no longer q_F^* because of $\lambda_4 > 0$. Binding $(A18)$, $(A19)$, and $(A21)$ give:

$$u^{\Delta}(q_H) = \frac{s}{(1-\beta)},$$

which implies $q_H^C = \overline{q_H} > q_H^*$. As in Case III, $\lambda_3 = 0$ implies that $q_L^C = q_L^*$. From the binding constraints, $p_H^C = u(\overline{q_H}, H) - \frac{s}{1-\beta} \left(< p_H^F\right)$ and $p_L^C = u(q_L^*, H) - \frac{s}{1-\beta} \left(\gtreqless p_L^F\right)$. As s increases, p_L^C decreases and eventually the constraint linked to λ_2 become no longer binding at $s = \overline{\overline{s}}$, where $\overline{\overline{s}} = (1-\beta)u^{\Delta}(\overline{\overline{q_H}})$.

Case V: $\lambda_1 > 0$, $\lambda_2 = 0$, $\lambda_3 = 0$ and $\lambda_4 > 0$. Solving $(A14)$ and $(A15)$ together with $\lambda_2 = \lambda_3 = 0$, we obtain $\lambda_1 = \varphi_H + \beta\varphi_L$ and $\lambda_4 = \beta\varphi_L$. Thus, $(A16)$ is rewritten as:

$$u_q(q_H, H) = c - \frac{\beta\varphi_L}{\varphi_H} u_q^{\Delta}(q_H),$$

which implies $q_H^C = \overline{\overline{q_H}} > q_H^F = q_H^*$. As in Case III and IV, $\lambda_3 = 0$ implies that $q_L^C = q_L^*$. From the binding constraints, we have $p_H^C = u(\overline{\overline{q_H}}, H) - \frac{s}{1-\beta} \left(< p_H^F\right)$ and $p_L^C = u(q_L^*, L) - u_q^{\Delta}(\overline{\overline{q_H}}) - \frac{s}{1-\beta} \left(\gtreqless p_L^F\right)$.

It follows that Case I and II belong to the regime of $s < \underline{s}$, Case III to the regime of $s \in [\underline{\underline{s}}, \underline{s})$, and Case IV and V to the regime of $s > \overline{s}$. ■

Proof of Proposition 3.

By Proposition 1, the firm's expected profit with $\Psi = F$ is:

$$\Pi^F = \varphi_H \left[u(q_H^*, H) - cq_H^* - u^{\Delta}(q_L^F)\right] + \varphi_L \left[u(q_L^F, L) - cq_L^F\right] - s.$$

Also, by Proposition 2, the firm's expected profit with $\Psi = C$ is:

$$\Pi^C = \begin{cases} \varphi_H \left[u(q_H^*, H) - cq_H^* - u^{\Delta}(\underline{\underline{q_L}})\right] + \beta\varphi_L \left[u(\underline{\underline{q_L}}, L) - c\underline{\underline{q_L}}\right] & \text{for } s < \underline{\underline{s}}, \\ \varphi_H \left[u(q_H^*, H) - cq_H^* - \frac{s}{1-\beta}\right] + \beta\varphi_L \left[u(\underline{q_L}, L) - c\underline{q_L}\right] & \text{for } s \in [\underline{\underline{s}}, \underline{s}), \\ \varphi_H \left[u(q_H^*, H) - cq_H^* - \frac{s}{1-\beta}\right] + \beta\varphi_L \left[u(q_L^*, L) - cq_L^*\right] & \text{for } s \in [\underline{s}, \overline{s}], \\ \varphi_H \left[u(\overline{q_H}, H) - c\overline{q_H} - \frac{s}{1-\beta}\right] + \beta\varphi_L \left[u(q_L^*, L) - cq_L^* - \frac{s}{1-\beta}\right] & \text{for } s \in (\overline{s}, \overline{\overline{s}}], \\ \varphi_H \left[u(\overline{\overline{q_H}}, H) - c\overline{\overline{q_H}} - \frac{s}{1-\beta}\right] + \beta\varphi_L \left[u(q_L^*, L) - cq_L^* - u_q^{\Delta}(\overline{\overline{q_H}}) - \frac{s}{1-\beta}\right] & \text{for } s > \overline{\overline{s}}. \end{cases}$$

Lemma 8 Π^C *is non-increasing and weakly concave in* s*.*

Proof. Applying the envelope theorem to the Lagrangian, we have $\frac{\partial \Pi^C}{\partial s} = -\frac{\lambda_1(s)}{1-\beta} \leq 0$. Now, we need to show

$$sign\left(\frac{\partial^2 \Pi^C}{\partial s^2}\right) = sign\left(-\frac{\partial \lambda_1(s)}{\partial s}\right) \leq 0.$$

Let us find $\lambda_1(s)$. From $(A14)$, $\lambda_1 = \varphi_H - \lambda_3 + \lambda_4$. Solving $(A15)$ and $(A17)$, we obtain $\lambda_3 = \beta\varphi_L \frac{u_q(q_L, L) - c}{u_q^{\Delta}(q_L)}$. Similarly, solving $(A14)$ and $(A16)$, we obtain $\lambda_4 = \max\left\{\varphi_H \frac{c - u_q(q_H, H)}{u_q^{\Delta}(q_H)}, 0\right\}$.

As a result,

$$\lambda_1(s) = \varphi_H - \beta\varphi_L \frac{u_q(q_L, L) - c}{u_q^\Delta(q_L)} + \max\left\{\varphi_H \frac{c - u_q(q_H, H)}{u_q^\Delta(q_H)}, 0\right\}.$$

It is immediate that $\frac{\partial\lambda_1(s)}{\partial s} = 0$ when $s < \underline{\underline{s}}$, $s \in [\underline{s}, \overline{s}]$, and $s > \overline{\overline{s}}$, because q_H^*, $\overline{\overline{q_H}}$, q_L^*, and $\overline{\overline{q_L}}$ is independent of s. In other words, in these three regimes, Π^C is linearly decreasing. However, when $s \in [\underline{\underline{s}}, \underline{s})$ or $s \in (\overline{s}, \overline{\overline{s}}]$, we have to investigate the sign of $\frac{\partial\lambda_1(s)}{\partial s}$. First, when $s \in [\underline{\underline{s}}, \underline{s})$, $\lambda_1(s) = -\beta\varphi_L \frac{u_q(\underline{q_L}, L) - c}{u_q^\Delta(\underline{q_L})}$. Thus, a simple calculation gives

$$\frac{\partial\lambda_1(s)}{\partial s} = -\beta\varphi_L \frac{\left[u_{qq}(\underline{q_L}, L)u_q^\Delta(\underline{q_L}) - u_{qq}^\Delta(\underline{q_L})(u_q(\underline{q_L}, L) - c)\right]}{u_q^\Delta(\underline{q_L})^2} \frac{\partial \underline{\underline{q_L}}}{\partial s} > 0.$$

The terms in the bracket is negative under well-accepted, $\frac{u_{qq}(q,L)}{u_q(q,L)} \leq \frac{u_{qq}(q,H)}{u_q(q,H)}$. By applying the implicit function theorem to $u^\Delta(\underline{q_L}) = \frac{s}{1-\beta}$, we have $\frac{\partial \underline{\underline{q_L}}}{\partial s} = \frac{1}{(1-\beta)u_q^\Delta(\underline{q_L})} > 0$. Likewise, it can be easily checked that for $s \in (\overline{s}, \overline{\overline{s}}]$, $\lambda_1(s) = \varphi_H \frac{c - u_q(q_H, H)}{u_q^\Delta(q_H)}$ is decreasing in s. ■

Π^F is linearly decreasing, while Π^C is concavely decreasing in s. Also, note that $\Pi^F(s = 0) > \Pi^C(s = 0)$, while $\Pi^F(s = \overline{\overline{s}}) > \Pi^C(s = \overline{\overline{s}})$. Thus, if we find any s such that $\Pi^F(s) < \Pi^C(s)$, there should exist s_l and s_h so that $\Pi^F(s) \geq \Pi^C(s)$ for $s \leq s_l$ and $s \geq s_h$, but $\Pi^F(s) < \Pi^C(s)$ for $s \in (s_l, s_h)$. Next, consider $\underline{s} \equiv (1-\beta)u^\Delta(q_L^*)$. Then, we have:

$$\begin{aligned}\Pi^C(s = \underline{s}) - \Pi^F(s = \underline{s}) &= (1-\beta)u^\Delta(q_L^*) + \beta\varphi_L\left[u(q_L^*, L) - cq_L^*\right] \\ &\quad -\varphi_H\left[u^\Delta(q_L^*) - u^\Delta(q_L^F)\right] - \varphi_L\left[u(q_L^F, L) - cq_L^F\right]\end{aligned}$$

The first two terms are positive, whereas the last two terms are negative. Given L, as H increases (hence Δ increases), both $u^\Delta(q_L^*)$ and $u^\Delta(q_L^F)$ increase. On the other hand, $u(q_L^F, L) - cq_L^F$ become smaller due to a greater downward distortion in q_L^F. Therefore:

$$\lim_{H\to\infty} \Pi^C(s = \underline{s}) - \Pi^F(s = \underline{s}) > 0,$$

which implies that $\Pi^F(s) < \Pi^C(s)$ for $s \in (s_l, s_h)$, where $s_l < \underline{s} < s_h$. ■

Proof of Proposition 4.

First, (PC_H) is implied by $(\widetilde{LC})$ and $(\widetilde{PC}_L)$:

$$\begin{aligned} u(q_H, H) - p_H - s &\geq \beta\left[u(q_L, H) - p_L\right] \\ &\geq \beta\left[u(q_L, L) - p_L\right] \geq 0 \end{aligned}$$

Thus, (PC_H) is not binding. In addition, we prove below that $(\widetilde{IC}_L)$ is not binding.

Lemma 9 *$(\widetilde{IC}_L)$ is never binding.*

Proof. $(\widetilde{LC})$ is written as $s \geq u(q_H, H) - p_H - \beta\left(u(q_L, H) - p_L\right)$. Similarly, $(\widetilde{IC}_L)$ is written as $s \geq u(q_H, L) - p_H - \beta\left[u(q_L, L) - p_L\right]$. Note that $(\widetilde{IC}_L)$ is always satisfied because $u^{\Delta}(q_H) > \beta u^{\Delta}(q_L)$ whether $(\widetilde{LC})$ is binding or not. ■

The Lagrangian of the firm's problem can be written as:

$$\begin{aligned} \mathcal{L} = \ &\varphi_H\left(p_H - cq_H\right) + \beta\varphi_L\left(p_L - cq_L\right) \\ &+\mu_1\left[u(q_H, H) - p_H - s - \beta\left(u(q_L, H) - p_L\right)\right] \\ &+\mu_2[u(q_L, L) - p_L] \\ &+\mu_3\left[u(q_H, H) - p_H - u(q_L, H) + p_L\right] \end{aligned}$$

The first order conditions are:

$$\frac{\partial \mathcal{L}}{\partial p_H} = \varphi_H - \mu_1 - \mu_3 = 0, \tag{5}$$

$$\frac{\partial \mathcal{L}}{\partial p_L} = \beta\varphi_L + \beta\mu_1 - \mu_2 + \mu_3 = 0, \tag{6}$$

$$\frac{\partial \mathcal{L}}{\partial q_H} = -\varphi_H c + \left(\mu_1 + \mu_3\right) u_q(q_H, H) = 0, \tag{7}$$

$$\frac{\partial \mathcal{L}}{\partial q_L} = -\beta\varphi_L c + \mu_2 u_q(q_L, L) - \left(\mu_1\beta + \mu_3\right) u_q(q_L, H) = 0, \tag{8}$$

$$\frac{\partial \mathcal{L}}{\partial \mu_1} = u(q_H, H) - p_H - s - \beta\left(u(q_L, H) - p_L\right) \geq 0, \quad \mu_1\frac{\partial \mathcal{L}}{\partial \mu_1} = 0, \tag{9}$$

$$\frac{\partial \mathcal{L}}{\partial \mu_2} = u(q_L, L) - p_L \geq 0, \quad \mu_2\frac{\partial \mathcal{L}}{\partial \mu_2} = 0, \tag{10}$$

$$\frac{\partial \mathcal{L}}{\partial \mu_3} = u(q_H, H) - p_H - u(q_L, H) + p_L \geq 0, \ \mu_3\frac{\partial \mathcal{L}}{\partial \mu_3} = 0, \tag{11}$$

Case I: $\mu_1 = 0, \mu_2 > 0$, and $\mu_3 > 0$. Suppose s is small enough. Then, the problem is reduced to a standard problem. We obtain $q_H^{\widetilde{C}} = q_H^*$ and $q_L^{\widetilde{C}} = \underline{\underline{q_L}}$ as in Case I in the full menu case. The non-binding constraint associated with μ_1 is written as:

$$s < u_q^{\Delta}(q_L^{\widetilde{C}}) - \beta u^{\Delta}(q_L^{\widetilde{C}}).$$

As s becomes large, the constraint will be binding at $s = \underline{\underline{s}}^{\tilde{c}}$, where $\underline{\underline{s}}^{\tilde{c}} \equiv u_q^{\Delta}(\underline{\underline{q_L}}) - \beta u^{\Delta}(\underline{\underline{q_L}})$.

Case II: $\mu_1 > 0, \mu_2 > 0$, and $\mu_3 > 0$. As s becomes $\underline{\underline{s}}^{\tilde{c}}$, the three constraints, (9), (10), and (11), simultaneously binding. Thus, in this case, $q_H^{\widetilde{C}} = q_H^*$, and $q_L^{\widetilde{C}} = \underline{q_L^{\widetilde{C}}}$ is characterized as

$$s = u_q^{\Delta}(\underline{q_L^{\widetilde{C}}}) - \beta u^{\Delta}(\underline{q_L^{\widetilde{C}}}).$$

This case is valid when $s \in [\underline{\underline{s}}^{\tilde{c}}, \underline{s}^{\tilde{c}})$, where $\underline{s}^{\tilde{c}} = u_q^{\Delta}(q_L^*) - \beta u^{\Delta}(q_L^*)$. If s becomes greater than $\underline{s}^{\tilde{c}}$, (IC_H), i.e., (11), becomes no longer binding.

Case III: $\mu_1 > 0, \mu_2 > 0$, and $\mu_3 = 0$. Next, when $s > \underline{s}^{\tilde{c}}$, only $(\widetilde{LC})$ and (PC_L) are binding. Therefore, we obtain $q_H^{\widetilde{C}} = q_H^*$ and $q_L^{\widetilde{C}} = q_L^*$ without any distortion. ■

Proof of Proposition 5.

By proposition 4, the firm's expected profit with $\Psi = \widetilde{C}$ is:

$$\Pi^{\widetilde{C}} = \begin{cases} \varphi_H \left[u(q_H^*, H) - cq_H^* - u^{\Delta}(\underline{\underline{q_L}})\right] + \beta\varphi_L \left[u(\underline{\underline{q_L}}, L) - c\underline{\underline{q_L}}\right] & \text{for } s < \underline{\underline{s}}^{\tilde{c}}, \\ \varphi_H \left[u(q_H^*, H) - cq_H^* - \left(\frac{u_q^{\Delta}(\underline{q_L^{\widetilde{C}}}) - s}{\beta}\right)\right] + \beta\varphi_L \left[u(\underline{q_L^{\widetilde{C}}}, L) - c\underline{q_L^{\widetilde{C}}}\right] & \text{for } s \in [\underline{\underline{s}}^{\tilde{c}}, \underline{s}^{\tilde{c}}), \\ \varphi_H \left[u(q_H^*, H) - cq_H^* - \beta u^{\Delta}(q_L^*) - s\right] + \beta\varphi_L \left[u(q_L^*, L) - cq_L^*\right] & \text{for } s \geq \underline{s}^{\tilde{c}}. \end{cases}$$

Like $\Pi^C(s)$, it can be shown that $\Pi^{\widetilde{C}}(s)$ is non-increasing and weakly concave in s. It is clear that $\Pi^C(s) = \Pi^{\widetilde{C}}(s)$ at $s = \min\{\underline{\underline{s}}, \underline{\underline{s}}^{\tilde{c}}\}$ and $s = \min\{\underline{s}, \underline{s}^{\tilde{c}}\}$. Also, note that $\Pi^{\widetilde{C}}(s)$ is decreasing in s more slowly than $\Pi^C(s)$ and $\Pi^F(s)$ for $s > \max\{\underline{s}, \underline{s}^{\tilde{c}}\}$. Thus, if we find $\Pi^{\widetilde{C}}(s) < \Pi^C(s)$ at a certain s, then there must exist $s_{\tilde{l}} < s_l$ such that $\Pi^{\widetilde{C}}(s) > \max\{\Pi^C(s), \Pi^F(s)\}$ for $s > s_{\tilde{l}}$. Below we provide a sufficient condition for it. We compare $\Pi^{\widetilde{C}}(s)$ and $\Pi^C(s)$ at $s = \underline{s}$ to find:

$$\Pi^{\widetilde{C}}(s = \underline{s}) - \Pi^C(s = \underline{s}) < 0 \Leftrightarrow u^{\Delta}(q_L^*) > u_q^{\Delta}(q_L^*).$$

We can obtain $u^{\Delta}(q_L^*) > u_q^{\Delta}(q_L^*)$ when q_L^* is large enough. Therefore, if c is small enough, the firm's choice of $\Psi = \widetilde{C}$ generates a greater profit when s is sufficiently large. ■

References

[1] Bar-Issac, H. et al. (2010), "Information Gathering and Marketing," *Journal of Economic and Management Strategy*, 19, 375–401.

[2] Brown, C. and Carpenter, G. S. (2000), "Why Is the Trivial Important? A Reason-Based Account for the Effects of Trivial Attributes on Choice," Journal of Marketing Research, 31, 339-350.

[3] Carlson, J. and McAfee, P. (1983), "Discrete Equilibrium Price Dispersion," *Journal of Political Economy*, 91, 480–493.

[4] Johnson, J. and Myatt, D. (2006), "On the Simple Economics of Advertising, Marketing and Product Design," *American Economic Review*, 96, 756–784.

[5] Jullien, B. (2000), "Participation Constraints in Adverse Selection Model," *Journal of Economic Theory*, 93, 1–47.

[6] Lewis, T. and Sappington, D. (1989), "Countervailing Incentives in Agency Problems," *Journal of Economic Theory*, 49, 294–313.

[7] Lewis, T. and Sappington, D. (1994), "Supplying Information to Facilitate Price Discrimination," *International Economic Review*, 35, 309–327.

[8] Maggi, G. and Rodriguez-Clare, A. (1995), "On Countervailing Incentives," *Journal of Economic Theory*, 66, 238–263.

[9] Maskin, E. and Riley, J. (1984), "Monopoly with Incomplete Information," *Rand Journal of Economics*, 15, 171–196.

[10] Mussa, M. and Rossen, S. (1978), "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

[11] Ottaviani, M. and Prat, A. (2001), "The Value of Public Information in Monopoly," *Econometrica*, 69, 1673–1683.

[12] Persico, N. (2000), "Information Acquisition in Auctions," *Econometrica*, 68, 135–148.

[13] Reinganum, J. (1979), "A Simple Model of Equilibrium Price Dispersion," *Journal of Political Economy*, 87, 851–858.

[14] Stiglitz, J. (1979), "Equilibrium in Product Markets with Imperfect Information," *American Economic Review*, 69, 339–345.

[15] Thompson, B. V. et al (2005), "Feature Fitigue: When Product Capabilities Become Too Much of a Good Thing," Journal of Marketing Research, 42, 431-442.