Adaptation and The Provider's Dilemma

Christina Aperjis

Social Computing Group, HP Labs, Palo Alto, CA 94304

Bernardo A. Huberman

Social Computing Group, HP Labs, Palo Alto, CA 94304

While attracting attention is one of the prime goals of content providers, the conversion of that attention into revenue is by no means obvious. Given that most users are used to consuming web content for free, a content provider faces a dilemma. Since the introduction of advertisements or subscription fees will be construed by users as an inconvenience which may deter them from using the website, what should the provider do in order to maximize revenues? We address this question through the lens of adaptation theory, which states that even though a change affects a person's utility initially, as time goes on people tend to adapt and become less aware of past changes. We establish that if the likelihood of attending to the provider is a log-convex function of the deviation of the total inconvenience from the reference point of a potential user, then it is always optimal for the provider to perform the increase in one step. Otherwise, the provider faces a tradeoff between achieving a higher revenue per user sooner and maximizing the number of users in the long term.

Key words: Internet monetization; online advertising; pricing; reference effects; adaptation

1.   Introduction

The goal of content providers is to turn attention to their websites into revenues that will at least offset their costs. But this is not an easy task, as even providers with established audiences often struggle to convert the attention they receive into profits. There are many ways of converting attention to revenue; charging subscription rates and presenting adverts are typical examples. Mixed strategies, where subscription fees and advertising are combined, have also been considered (Baye and Morgan 2000, Prasad et al. 2003, Kumar and Sethi 2009). But all these strategies carry a price, for while some people perceive the associated costs as an inconvenience to be tolerated in exchange for the value obtained, others see them as a nuisance that makes users leave the website and deters other potential users from joining. This issue has been especially acute with the advent of increasingly intrusive "rich media" advertising formats (Godes et al. 2009).

Given that a provider can increase his revenue by imposing some inconvenience to users while risking losing some of the attention paid to his content, how steeply and for how long should he increase this inconvenience in order to maximize revenue? We address this question through the lens of adaptation theory, which states that even though a change affects a person's happiness in the short term, in the course of time people tend to adapt and become less aware of past changes (Frederick and Loewenstein 1999, Frey and Stutzer 2002). Furthermore, as a number of empirical studies show, gradual changes and spikes in utility have rather different effects on adaptation levels: whereas sudden changes are noticed and evaluated, a very slow gradual change will drag the adaptation level along with it and at times may not even be detected (Kahneman and Thaler 1991). One might thus expect that it is better for the provider to increase inconvenience gradually over time, because that would result in more users in the long term. However, we show that this is not always true: in certain cases, which we characterize, it is optimal to increase inconvenience once.

We treat the dynamics of the adaptation process in settings without competition where multiple changes can occur over time. Each user has a reference point which depends on past inconvenience levels at the website; the reference points of non-users are equal to zero.

Our analysis is based on the probability of attending to the provider as a function of the difference between inconvenience and reference point; a function that can be measured in real settings (e.g., with A/B testing). The shape of this function characterizes the strategy that maximizes the provider's profit. We find that if it is log-convex (i.e., the logarithm of the function is convex), then it is always optimal for the provider to increase inconvenience only one time. On the other hand, if the probability of attending is not log-convex, it is usually optimal to increase inconvenience in multiple stages. The provider then faces a tradeoff between achieving a higher profit per user sooner, and maximizing the number of users in the long term. We study the provider's profit optimization problem for this setting.

Adaptation theory allows us to consider how users react over time to an introduced inconvenience. A number of papers consider adaptation in the context of repeat-purchase markets and characterize optimal dynamic pricing policies (Kopalle et al. 1996, Fibich et al. 2003, Popescu and Wu 2007, Nasiry and Popescu 2010). In these papers, a firm (usually a monopolist) is facing consumers whose purchase decisions are influenced by past prices through reference price effects. The demand in a given period is assumed to be a function of the current price and the reference price (but does not depend on the number of people that purchased the product in the previous period).

Our paper complements the literature on dynamic pricing with reference effects by allowing for users that joined at different points in time in the past to have different reference points (because the time that has elapsed since a user joined influences how much his reference point has adapted). Moreover, we assume that non-users can have yet a different reference point, because they may have not adapted to the total inconvenience of the website.[1] As a result, the demand (number of users) in a given period is influenced by the demand in the previous period. Furthermore, we study revenue optimization for a content provider on the web and consider a general framework that applies for any type of inconvenience to the user that generates profit for the provider; subscription fees and advertisements are typical examples. More importantly, in contrast to the prior literature on dynamic pricing with reference effects where it is optimal to introduce changes gradually, we identify two qualitatively different regimes: in one it is optimal to introduce the full inconvenience at once, whereas in the other it is better to gradually increase inconvenience over time.

A number of experimental and empirical studies have focused on the formation of reference points (surveys are provided by Kalyanaram and Winer 1995, Mazumdar et al. 2005). In these studies, the inconvenience is the price of a product, and thus the reference point is a reference price. Even though the role of historic prices in forming price expectations is supported in many of these studies, there has not been sufficient evidence for any specific model on how consumers update their reference prices. Our main result on log-convexity holds under very general assumptions on how the reference point adapts over time, which includes exponential smoothing (Kopalle et al. 1996, Fibich et al. 2003, Popescu and Wu 2007), peak-end anchoring (Nasiry and Popescu 2010) and linear adaptation (Chen and Rao 2002) as special cases.

The paper is structured as follows. Section 2 introduces the model. In Section 3 we study the provider's profit optimization problem and show our main result. Section 4 briefly discusses more general probability models, where the probability of attending to the provider is a function of both the deviation from the reference point and the total inconvenience. Section 5 concludes. All the proofs are provided in the Appendix.

---

[1] This is similar to a distinction made in the marketing literature between consumers that are loyal to a brand and non-loyal consumers (e.g., Krishnamurthi et al. 1992).

2. The Model

We consider a website owner (the provider) that wishes to maximize his profits though advertisements or subscription costs. Each of these profit generating processes imposes some "inconvenience" to users. In particular, it is often time-consuming for users to view advertisements, whereas targeted advertising may violate a consumer's privacy (Goldfarb and Tucker 2011). On the other hand, subscription fees are clearly undesirable to users. In what follows, we abstract from the specific profit creating process and talk about increasing inconvenience to the users instead.

We consider discrete time $t = 1, 2, 3, ...$ and assume that in each period the provider has the option to increase inconvenience. We denote the total inconvenience level (e.g., advertisement level, subscription cost) at time t by X_t. We define $x_t \equiv X_t - X_{t-1}$ to be the increase in inconvenience at time t. We make the following assumptions, motivated by the fact that websites often start without imposing any inconvenience to users in order to attract them and then increase inconvenience over time.

ASSUMPTION 1. *(i)* $X_0 = 0$, *that is, initially there is no inconvenience.*
(ii) $x_t \geq 0$, *that is, the provider does not decrease the inconvenience at any point in time.*

Assumption 1 implies that $X_t = \sum_{j=1}^{t} x_j$.

In each period, each potential user has a personal reference point and as a result ends up using the website with some probability which depends on how much the current level of inconvenience exceeds his reference point. The reference points of users gradually adapt over time. Different people may have different reference points. A key aspect of our approach is that users have a different reference point than non-users. Moreover, we allow a user's reference point to depend on when he joined.

2.1. Consumers

Let N_0 be the number of users at time 0, i.e., before any inconvenience is introduced. We assume that S_t is the number of non-users that consider becoming users at time t.[2] For simplicity, we assume that each S_i is deterministic; however, our analysis would also go through if each S_i were a random variable. Each person that uses the website at time $t - 1$ will also consider to use the website at time t. We refer to users that joined at time t as type t users. Moreover, we refer to users that were there since time 0 as type 0 users.

We assume that a person with reference point r_t that considers using the website at time t will end up using it with probability $p(X_t - r_t)$, that is, we assume that this probability only depends on the difference between the current inconvenience and the reference point. (More general probability models are discussed in Section 4.) We assume that p is a decreasing function: the larger the difference between total inconvenience and the reference point, the smaller the probability of using the website.

We now consider the reference points of non-users and users.

Non-users. Let r_t^t be the reference point of a non-user that considers joining at time t. We assume that $r_t^t = 0$ (for all $t \geq 1$), that is, *non-users have a degenerate reference point* because they have not adapted to the inconvenience that users are experiencing on the website.[3] Thus, a non-user that considers to join the website at time t ends up joining with probability $p(X_t)$.

---

[2] In this paper, we assume that the sequence $\{S_i, i = 1, 2, ...\}$ is exogenous. Alternatively, the number of non-users that consider to join at time t could depend on the number of users at time $t - 1$. For instance, S_t and the number of users at time $t - 1$ would be positively correlated in the case of word of mouth or herding and negatively correlated in the case of congestion effects (but we do not consider these effects here).

[3] The assumption $r_t^t = 0$ is not essential; more generally, our results hold if $r_t^t \leq X_t$. We note that in a setting with competition, the reference point could also depend on the average inconvenience in the market or on the inconvenience on other websites that the particular person has used. The reference brand (studied by Kopalle et al. 1996) is a special case. However, we do not consider competition in this paper.

Users. We denote the reference point of a type i user at time t by r_t^i (where $i < t$). At time t, such a user will continue using the website with probability $p(X_t - r_t^i)$. According to adaptation theory, even though a change initially affects a person's happiness, as time goes on people tend to adapt and become less aware of past changes. Here we do not consider a specific model for how the reference point adapts over time. We only assume that $r_t^i \leq X_{i-1}$, which is satisfied by all reference point formation mechanisms that depend on historic inconvenience levels (because of Assumption 1). For instance, exponential smoothing (the most commonly used reference point formation mechanism) in our context would imply that for some $\alpha \in [0,1)$, $r_t^i = \alpha \cdot r_{t-1}^i + (1-\alpha)X_{t-1}$ for $t > i$. Then, since $r_i^i = 0$ and $X_j \leq X_{j+1}$, we have that $r_t^i \leq X_{i-1}$. Other reference point formation mechanisms, such as peak-end anchoring (Nasiry and Popescu 2010) and linear adaptation (Chen and Rao 2002) also satisfy $r_t^i \leq X_{i-1}$ in our setting.

We note that complete adaptation is a special case: if existing users have completely adapted to the level of inconvenience in the previous period (i.e., $r_t = X_{t-1}$), the probability that a user continues using the website at time t is equal to $p(X_t - X_{t-1}) = p(x_t)$.

We next summarize our assumptions on the reference points.

ASSUMPTION 2.   *(i)* $r_t^i \leq X_{t-1}$ *for* $t > i$
(ii) $r_t^t = 0$, *that is, non-users have a degenerate reference point*

Our analysis and results rely heavily on the function p. We next describe a utility model from which p may arise; we point out, however, that because our results are stated in terms of p they hold even in the absence of a utility model.

Additive Random Utility Model and Prospect Theory We consider the user's *experienced utility*, that is, the hedonic experience associated with the use of a website (see Kahneman and Thaler 2006) and take a prospect theory approach (Kahneman and Tversky 1979). Prospect theory proposes that preferences are defined by the deviation from a reference point rather than by the final state of the outcome: positive deviations are coded as gains and negative deviations as losses. In our setting, we are only interested in negative deviations where inconvenience increases (because of Assumption 1).

Let r_t be a person's reference point for inconvenience at time t and let X_t be the total level of inconvenience that is currently experienced on the website. According to prospect theory, the person's utility from using the website is going to be a function of the deviation from the reference point (which is equal to $X_t - r_t$) and will not directly depend on the final state (which is X_t). We write this utility as $v(-(X_t - r_t))$, where v denotes the value function of prospect theory. The standard assumptions on v are that it is increasing throughout its domain, convex on $(-\infty, 0)$ (which is the relevant domain for our analysis), concave on $(0, \infty)$ and $v(0) = 0$ (Kahneman and Tversky 1979).

In order to model heterogeneity in the population of consumers, we perturb the quantity $v(-(X_t - r_t))$ with a random term Y, which represents the benefit to the consumer from using the website. We thus assume that the utility from using the website at time t for the particular consumer is $v(-(X_t - r_t)) + Y$. Moreover, we assume that the utility from not using the website is equal to zero. Thus, the person will use the website at time t with probability

$$\mathbb{P}[v(-(X_t - r_t)) + Y > 0] = 1 - \mathbb{P}[Y < -v(-(X_t - r_t))].$$

Let F be the cumulative distribution function of Y, and define

$$p(x) \equiv 1 - F(-v(-x)). \tag{1}$$

According to the Additive Random Utility Model (ARUM)[4] that we just introduced, a person with reference point r_t uses the website (at time t) with probability $p(X_t - r_t)$.

---

[4] We note that ARUMs are the standard way to model discrete choice in economics (e.g., Cameron and Trivedi 2005).

We can now relate this model to adaptation. According to adaptation theory, even though a change initially affects a person's happiness, as time goes on people tend to adapt and become less aware of past changes. In the context of our theory, an increase in inconvenience by an amount x initially decreases a user's utility. However, as time goes by the user's reference point gradually adapts and, as a result, his experienced utility gradually increases if no additional inconvenience is experienced.

2.2. Provider

The provider can increase inconvenience at the beginning of every period. The provider wishes to maximize his profit, which at any given point in time is an increasing function of both the number of users and the current inconvenience level. We denote the provider's profit per user (per period) from an inconvenience of x by $\pi(x)$. We assume that π is an increasing function: the higher the inconvenience to the user, the higher the profit to the provider. If this were not true, there would be no conflict of interest: the provider would decrease the inconvenience to make both himself and the users better off. Furthermore, we assume that the provider discounts future payments in that he prefers to get profit sooner than later. We denote the provider's discount factor by δ.

3. Profit Maximization

In this section, we consider the provider's profit optimization problem. We study how steeply and for how long the provider should increase inconvenience in order to maximize his infinite horizon expected discounted profit.

The provider solves the following problem:

$$\text{maximize } \Pi(\vec{x}) \equiv N_0 \sum_{t=0}^{\infty} \delta^t \prod_{j=1}^{t} p(X_j - r_j^0) \cdot \pi(X_t) + \sum_{i=1}^{\infty} S_i \cdot p(X_i) \sum_{t=i}^{\infty} \delta^t \prod_{j=i+1}^{t} p(X_j - r_j^i) \cdot \pi(X_t)$$

$$\text{subject to } X_t = \sum_{j=1}^{t} x_j, \text{ for } t \geq 1$$

$$x_t \geq 0, \text{ for } t \geq 1$$

In particular, $\Pi(\vec{x})$ denotes the provider's infinite horizon profit from introducing inconvenience $\vec{x} = (x_1, x_2, x_3, ...)$, where x_t is the inconvenience he introduces in period t. The first term gives the expected profit from the people using the website at time 0 (the established user base) and the second term gives the expected profit from the users that join later. Each user in the established user base (which consists of N_0 users) continues using the website at time t with probability $\prod_{j=1}^{t} p(X_j - r_j^0)$. On the other hand, each of the S_i non-users that consider joining at time i ends up joining with probability $p(X_i)$. Moreover, at time j, a current type i user continues using the website with probability $p(X_j - r_j^i)$. Thus, a non-user that considers joining at time i is a user at time t with probability $p(X_i) \prod_{j=i+1}^{t} p(X_j - r_j^i)$. The profit per user at time t is equal to $\pi(X_t)$. The profit that the provider will receive at time t is discounted by δ^t.

In the following section we show that if the probability p is log-convex, then it is optimal to introduce the full inconvenience at once. We then consider the problem of maximizing $\Pi(\vec{x})$ for a general probability function p when the profit π is log-concave (under some simplifying assumptions).

3.1. Log-convex probability

In this section, we consider the case of a log-convex p. A function is log-convex if its logarithm is convex. For instance, this is the case if $p(x) = 1/(1+x)^k$ with $k > 0$ or $p(x) = e^{-x^k}$ with $k \in (0, 1)$.

The setting of a log-convex p is of particular interest for profit maximization, because in that case it is optimal for the provider to increase inconvenience once (and not gradually over time). This is the context of the following proposition.
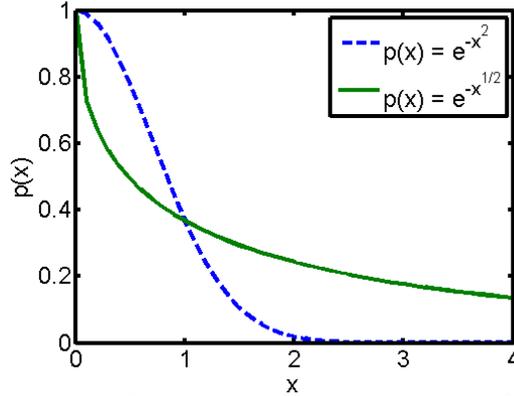
Figure 1 **A comparison between a function that is log-convex and a function that is not log-convex.**

PROPOSITION 1. *Let*

$$x^* \in \arg\max_{x \geq 0}\{p(x) \cdot \pi(x)\}.$$

If p is log-convex, then $(x^*, 0, 0, ...)$ is a maximizer of $\Pi(x_1, x_2, x_3, ...)$.

The proof of Proposition 1 shows that for any $\vec{x} = (x_1, x_2, x_3, ...)$ with $\sum_{i \geq 2} x_i > 0$, it is possible to find some x_1' such that $\Pi(x_1', 0, 0, ...) \geq \Pi(\vec{x})$. In words, for any trajectory of increase in inconvenience, it is possible to find a solution with a single increase that is at least as good in terms of discounted expected profits. Thus, the problem of maximizing $\Pi(\vec{x})$ reduces to maximizing $p(x) \cdot \pi(x)$. This is a significant simplification on the problem: the provider does not need to know how people adapt over time nor how many non-users will consider joining at each period in the future.

One might expect that in the presence of adaptation, it would be better for the provider to increase inconvenience gradually over time, because that would give people more time to adapt to changes and presumably result in more users in the long term. Proposition 1 shows that our intuition is led astray in the case of a log-convex p: if p is log-convex, it is optimal to increase inconvenience once — *despite adaptation effects*.

To get some intuition for this, consider Figure 1 which shows the log-convex function $e^{-x^{1/2}}$ and the function e^{-x^2} (which is not log-convex). We observe that for small deviations from the reference point, a consumer is less likely to use the website when his behavior is described by the log-convex function. On the other hand, for large deviations, a consumer is more likely to use the website when his behavior is described by the log-convex function compared to the the case that his behavior is not described by a log-convex. This suggests that under a log-convex function it is better to make one large change, whereas in other cases it is better to make many small changes over time. We note that this is *not* a result of selection, because the function p does not change over time.

We close this section by discussing what assumptions of the ARUM, which was introduced in Section 2, give rise to a log-convex p. Recall that according to that ARUM, p is given by (1), where F is a cumulative probability distribution and v is the value function of prospect theory. The following lemma shows that p is log-convex if the reliability function (i.e., the complementary cumulative distribution function) is log-convex and v is convex.

LEMMA 1. *Suppose that (1) holds and F is twice-differentiable. If $1 - F(x)$ is log-convex and $v''(x) \geq 0$ for $x \in (-\infty, 0)$, then p is log-convex on $[0, \infty)$.*

According to prospect theory, v is convex on $(-\infty, 0)$ (Kahneman and Tversky 1979). Thus, if the reliability function $1 - F$ is log-convex, then p is log-convex. The following lemma gives a

condition in terms of the density and cumulative distribution functions under which the reliability function is log-convex.

LEMMA 2. *Consider a twice-differentiable cumulative probability function F and let f be the corresponding density. $1 - F(x)$ is log-convex if and only if*

$$-f'(x) \cdot (1 - F(x)) \geq f(x)^2$$

for all x.

Lemma 2 implies that a decreasing density distribution function is a necessary condition for log-convexity of the reliability function and p. In words, a randomly chosen consumer should be more likely to derive a small benefit from the website. A number of distributions have log-convex reliability functions (Bagnoli and Bergstrom 2005). For instance, this is the case for the exponential distribution, the Pareto distribution, the Weibull distribution when the shape parameter is smaller the one and the Gamma distribution when the shape parameter is smaller the one. Thus, if the random variable Y is drawn from one of these distributions, it is optimal to introduce the full inconvenience at once.

On the other hand, the function p may not be log-convex if the reliability function is not log-convex (which is the case for the uniform and the normal distribution). Then, the provider's profit optimization problem will not be equivalent to maximizing $p(x) \cdot \pi(x)$. This case is studied in the following section.

3.2.   Log-concave profit per user

In this section, we assume that the profit function π is log-concave.[5] A function is log-concave if its logarithm is concave. All concave and linear functions are log-concave, but there also exist convex functions that are log-concave.

Concavity and linearity are reasonable assumptions for a profit function, because such functions exhibit constant or decreasing marginal returns. Because of the generality of log-concavity, our results apply to a variety of situations. When the inconvenience is generated by a subscription fee, then the profit per user is equal to the subscription fee itself, and thus $\pi(x) = x$. On the other hand, when the source of inconvenience is advertising, then π can model various pricing schemes for online advertising (e.g., pricing per impression, pricing per click, and pricing per acquisition). Moreover, the price-per-impression and the price-per-click could either be exogenously defined or depend on the total number of advertisements on the website.

In this section, we consider profit maximization when π is log-concave under three simplifying assumptions. First, we assume that existing users completely adapt to changes in one period, that is, $r_t^i = X_{t-1}$ for $i < t$. Even though this assumption is certainly more restrictive than Assumption 2 (i), it is not unreasonable (e.g., see the arguments by Krishnamurthi et al. 1992).[6] Second, we assume no new users arrive over time, that is, $S_i = 0$ for $i \geq 1$. This case is of special interest, because it models an established provider with an existing user base whose main concern is to convert the attention he is already getting into profit (and does not focus on growth) — a concern that many established providers face (e.g., The New York Times, The Times, YouTube, FaceBook, Twitter). Lastly, we assume that the magnitude of all increases is the same.[7]

---

[5] Note that in this section it is the profit function π that is assumed to be log-concave, whereas in the previous section it was the probability function p that was assumed to be log-convex.

[6] It is possible to extend the results of this section for the case that users completely adapt in several periods (instead of just in one) under the assumption that the provider does not further increase inconvenience until users have completely adapted to the current level.

[7] More general cases can be solved numerically.
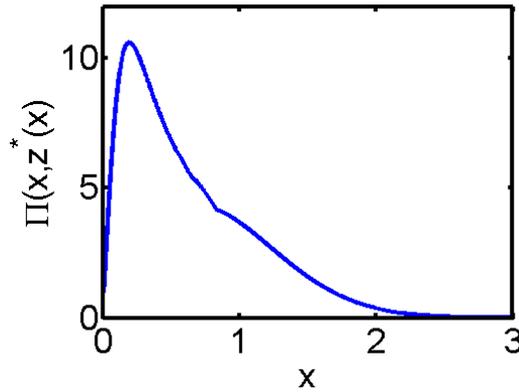
Figure 2 $\Pi(x, z^*(x))$ **for Example 1. It is maximized at** 0.195**, which implies that the revenue is maximized if the inconvenience increases** $z^*(0.195) = 26$ **times by an amount of** 0.195**.**

Then, the maximization of $\Pi(\vec{x})$ reduces to a two-dimensional problem, since we need to identify the magnitude of each increase and the number of times that it is introduced. At the risk of overloading notation, we denote by $\Pi(x, z)$ the discounted expected profit when inconvenience is increased z times by x, that is, $x_i = x$ for $i = 1, 2, ..., z$ and $x_i = 0$ for $i > z$. We observe that under complete adaptation and no new users,

$$\Pi(x, z) = N_0 \sum_{t=1}^{\infty} \delta^t p(x)^{\min\{t, z\}} p(0)^{(t-z)^+} \pi(\min\{t, z\} x), \tag{2}$$

where $y^+ \equiv \max(y, 0)$ denotes the positive part of y. For a fixed x, let $z^*(x)$ be the maximizer of $\Pi(x, z)$. The following lemma shows that the two-dimensional problem of maximizing $\Pi(x, z)$ can be reduced to a one-dimensional problem by identifying $z^*(x)$.

LEMMA 3. *If π is log-concave, then for a fixed x,*

$$z^*(x) = \min \left\{ z \in \mathbb{N} : p(0) \frac{\pi(x \cdot z)}{\pi(x \cdot (z+1))} \geq p(x) \right\}.$$

We now provide some intuition for this result. After the z-th change is introduced, the provider gets $\pi(x \cdot z)$ from each user. Increasing the inconvenience by x one more time will result in a profit of $\pi(x \cdot (z+1))$ from each remaining user and each user will stay with probability $p(x)$ (whereas each user will stay with probability $p(0)$ if there is no increase). Thus the change is worthwhile if and only if

$$p(0) \cdot \pi(x \cdot z) \leq p(x) \cdot \pi(x \cdot (z+1)).$$

Because π is log-concave, the ratio

$$\frac{\pi(x \cdot z)}{\pi(x \cdot (z+1))}$$

is increasing in z, which implies that it is never profitable to increase z above $z^*(x)$. Note that $z^*(x)$ does not depend on the discount factor δ.

Lemma 3 does not make any assumptions about p. It applies for any p, whether it is log-convex or not. However, since we already know (from Proposition 1) how to maximize $\Pi(\vec{x})$ when p is log-convex, Lemma 3 will be useful when p is not log-convex. The following example applies Lemma 3 to maximize the provider's profit for an instance of the problem.

EXAMPLE 1. Suppose $\delta = 0.9$, $p(x) = e^{-x^2}$ and $\pi(x) = x$. We plot $\Pi(x, z^*(x))$ in Figure 2. At the optimal solution $(x, z) = (0.195, 26)$, that is, the revenue is maximized if the inconvenience increases 26 times by an amount of 0.195. Thus, if $p(x)$ represents the probability of staying when the subscription fee is increased by x dollars, then it is optimal to increase the subscription fee by an amount of \$0.20 for a total of 26 times until reaching a final subscription fee of approximately \$5.20.

4. Other Probability Models

Up to now, we have assumed that the probability of using the website at a given point in time depends solely on the deviation from the reference point (that is, the difference between the total inconvenience and the reference point). However, more generally, the probability could depend on both the deviation from the reference point and the total inconvenience introduced so far. We write $q(X_t - r_t, X_t)$ for the probability in this more general case. One special case is of course the case $q(X_t - r_t, X_t) \equiv p(X_t - r_t)$, which has been studied in the previous sections.

We now consider the case that $q(x, X)$ is decreasing in X, that is, for a fixed deviation from the reference point, the probability of using the website is decreasing in the total inconvenience. Proposition 1 can be generalized for this setting. In particular, for any function $q(x, X)$ that is non-increasing in X and log-convex in x, the problem of maximizing the provider's profit reduces to maximizing $q(x, 0) \cdot \pi(x)$. The optimal strategy is then to introduce inconvenience of magnitude $x^* \in \arg\max\{q(x, 0) \cdot \pi(x)\}$ right away (through a single increase).

Lemma 3 can also be generalized for specific forms of $q(x, X)$ that are decreasing in X (under the assumptions of Section 3.2). In particular, equation (2) can be generalized to

$$\Pi(x, z) = N_0 \sum_{t=1}^{\infty} \delta^t \left(\prod_{j=0}^{\min\{t,z\}-1} q(x, j \cdot x) \right) (q(0, z \cdot x))^{(t-z)^+} \pi(\min\{t, z\}x).$$

If π is log-concave and $q(x, j \cdot x)/q(0, j \cdot x)$ is decreasing in j, then for a fixed x, $\Pi(x, z)$ is maximized at

$$z^*(x) = \min\left\{ z \in \mathbb{N} : q(0, z \cdot x) \frac{\pi(x \cdot z)}{\pi(x \cdot (z+1))} \geq q(x, z \cdot x) \right\}.$$

Additive Random Utility Model Generalization

A function $q(x, X)$ that is decreasing in X can arise from an Additive Random Utility Model if we consider a generalization of prospect theory. In their seminal paper on prospect theory, Kahneman and Tversky (1979) proposed that preferences are defined by the deviation from a reference point rather than by the final state of the outcome. Later, Kőszegi and Rabin (2006) developed a general model that includes both final outcome utility and gain-loss utility.

We can apply this general model to our context by assuming that the final outcome cost from a total inconvenience of X is $c(X)$, where c is an increasing function. Then, using an ARUM — similarly to Section 2, we find that

$$q(x, X) = 1 - F(c(X) - v(-x)).$$

This is clearly a decreasing function of X, formalizing the intuition that *the greater the total level of inconvenience the less likely a potential consumer is to use the website.*

5. Discussion

This paper studies profit maximization from the point of view of a content provider through the lens of adaptation theory. The provider can increase profits by imposing some inconvenience to users while risking to lose some of the current users as well as some potential future users. Our approach is very general in that it can be applied for any profit generating process that imposes inconvenience to the users (e.g., advertisements, subscription fees). Our results complement the existing work on dynamic pricing with reference effects by identifying a regime in which it is optimal for the provider to introduce the full inconvenience at once.

Our analysis is based on the function p that represents the probability that a person uses the website in a given period as a function of how much inconvenience deviates from his reference point. We provide a utility model from which p may arise; however, knowledge of the utility model is not essential for applying the results. In particular, the provider can directly use p to find the optimal strategy that maximizes his profit (if he has a reasonable model for how people update their reference points over time). We have shown that if p is log-convex, then it is optimal to increase inconvenience once; therefore, in this case the provider does not need to know how users would update their reference points after an increase in inconvenience nor how many non-users will consider joining in the future.

The provider can use A/B testing to estimate p. For instance, to get an estimate of $p(\hat{x})$ for some value $\hat{x}$, the provider can impose this inconvenience to some users and measure the percentage of these users that continue using the website. The provider should only use a small percentage of users to estimate p. Once he has a good estimate for p through which he can compute the optimal way to introduce the inconvenience, then he can introduce the optimal inconvenience for all users. We note that p can also be estimated using information from past experience and surveys.

An interesting open question is whether the type of inconvenience typically determines the form of p — in terms of whether it is log-convex — across websites. In that case, our results could provide insights on the optimal strategies for certain types of inconvenience. For instance, if it turns out that p is not log-convex in the case of advertising, that would imply that providers should increase the amount of ads slowly over time. On the other hand, if p is typically log-convex in the case of subscription fees, that would imply that a content provider enacting a paywall should increase the subscription cost once.

References

Bagnoli, Mark, Ted Bergstrom. 2005. Log-concave probability and its applications. *Economic Theory* **26**(2) 445–469.

Baye, M. R., J. Morgan. 2000. A simple model of advertising and subscription fees. *Economics Letters* **69** 345–351.

Cameron, A. Colin, Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge University Press.

Chen, Haipeng (Allan), Akshay R. Rao. 2002. Close encounters of two kinds: False alarms and dashed hopes. *Marketing Science* **21**(2) 178–196.

Fibich, Gadi, Arieh Gavious, Oded Lowengart. 2003. Explicit solutions of optimization models and differential games with nonsmooth (asymmetric) reference-price effects. *Operations Research* **51**(5) 721–734.

Frederick, Shane, George Loewenstein. 1999. Hedonic adaptation. Daniel Kahneman, Ed Diener, eds., *Well-being: The foundations of hedonic psychology*. 302–329.

Frey, Bruno S., Alois Stutzer. 2002. What can economists learn from happiness research? *Journal of Economic Literature* **40**(2) 402–435.

Godes, David, Elie Ofek, Miklos Sarvary. 2009. Content vs. advertising: The impact of competition on media firm strategy. *Marketing Science* **28**(1) 20–35.

Goldfarb, Avi, Catherine E. Tucker. 2011. Privacy regulation and online advertising **57**(1) 57–71.

Kahneman, Daniel, Richard Thaler. 1991. Economic analysis and the psychology of utility: Applications to compensation policy. *The American Economic Review* **81**(2) 341–346.

Kahneman, Daniel, Richard H. Thaler. 2006. Anomalies: Utility maximization and experienced utility. *Journal of Economic Perspectives* **20**(1) 221–234.

Kahneman, Daniel, Amos Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47**(2) 263–292.

Kalyanaram, Gurumurthy, Russell S. Winer. 1995. Empirical generalizations from reference price research. *Marketing Science* **14**(3) 161–169.

Kőszegi, Botond, Matthew Rabin. 2006. A model of reference-dependent preferences **121**(4) 1133–1165.

Kopalle, Praveen K., Ambar G. Rao, Joao L. Assuncao. 1996. Asymmetric reference price effects and dynamic pricing policies. *Marketing Science* **15**(1) 60–85.

Krishnamurthi, Lakshman, Tridib Mazumdar, S. P. Raj. 1992. Asymmetric response to price in consumer brand choice and purchase quantity decisions. *Journal of Consumer Research* **19** 387–400.

Kumar, Subodha, Suresh P. Sethi. 2009. Dynamic pricing and advertising for web content providers. *European Journal of Operational Research* **197** 924–944.

Mazumdar, Tridib, S.P. Raj, Indrajit Sinha. 2005. Reference price research: Review and propositions. *Journal of Marketing* **69** 84–102.

Nasiry, Javad, Ioana Popescu. 2010. Dynamic pricing with loss averse consumers and peak-end anchoring. *INSEAD Working Paper No. 2010/102/DS* .

Popescu, Ioana, Yaozhong Wu. 2007. Dynamic pricing strategies with reference effects. *Operations Research* **55**(3) 413–429.

Prasad, Ashutosh, Vijay Mahajan, Bart Bronnenberg. 2003. Advertising versus pay-per-view in electronic media. *Intern. J. of Research in Marketing* **20** 13–30.

Appendix

Proof of Proposition 1: Suppose that some vector $(x_1, x_2, x_3, ...)$ with $\sum_{j \geq 2} x_j > 0$ is strictly better than $(x^*, 0, 0, ...)$. We will contradict this by showing that

$$\Pi(X_z, 0, 0, ...) \geq \Pi(x_1, x_2, x_3, ...)$$

for some z (where $X_z = \sum_{j=1}^{z} x_j$). This will imply that

$$\Pi(x_1, x_2, x_3, ...) \leq \Pi(X_z, 0, 0, ...) \leq \Pi(x^*, 0, 0, ...).$$

The log-convexity of p implies that if $y \geq 0$,

$$\log(p(x + y)) - \log(p(y)) \geq \log(p(x)) - \log(p(0)),$$

or equivalently,

$$p(x) \cdot p(y) \leq p(0) \cdot p(x + y). \tag{3}$$

Let $\gamma_{i,t}(x_1, x_2, x_3, ...)$ be the expected profit in period t from a type i user. Then

$$\Pi(x_1, x_2, x_3, ...) = \sum_{i=0}^{\infty} S_i \sum_{t=i}^{\infty} \delta^t \gamma_{i,t}(x_1, x_2, x_3, ...),$$

and

$$\gamma_{i,t}(x_1, x_2, x_3, ...) = \begin{cases} \prod_{j=1}^{t} p(X_j - r_j^0) \cdot \pi(X_t) & \text{if } 0 = i \leq t \\ p(X_i) \cdot \prod_{j=i+1}^{t} p(X_j - r_j^i) \cdot \pi(X_t) & \text{if } 1 \leq i \leq t \\ 0 & \text{if } i > t \end{cases}$$

(3) implies that

$$\prod_{j=i+1}^{t} p(X_j - r_j^i) \le p\left(\sum_{j=i+1}^{t}(X_j - r_j^i)\right) \cdot p(0)^{t-i-1} \le p\left(\sum_{j=i+1}^{t} x_j\right) \cdot p(0)^{t-i-1} = p(X_t - X_i) \cdot p(0)^{t-i-1},$$

where the second inequality follows from Assumption 2. This further implies that

$$p(X_i) \cdot \prod_{j=i+1}^{t} p(X_j - r_j^i) \le p(X_t) \cdot p(0)^{t-i}.$$

Let

$$z \in \arg\max_t \{p(X_t) \cdot \pi(X_t)\}.$$

Then

$$\gamma_{0,t}(x_1, x_2, x_3, ...) = \prod_{j=1}^{t} p(X_j - r_j^0) \cdot \pi(X_t) \le p(X_t) \cdot \pi(X_t) \cdot p(0)^{t-1} \le p(X_z) \cdot p(0)^{t-1} \pi(X_t) = \gamma_{0,t}(X_z, 0, 0, ...),$$

and for $1 \le i \le t$,

$$\gamma_{i,t}(x_1, x_2, x_3, ...) = p(X_i) \prod_{j=i+1}^{t} p(X_j - r_j^i) \pi(X_t) \le p(0)^{t-i} \cdot p(X_z) \cdot \pi(X_z) = \gamma_{i,t}(X_z, 0, 0, ...).$$

Thus,

$$\Pi(x_1, x_2, x_3, ...) = \sum_{i=0}^{\infty} S_i \sum_{t=i}^{\infty} \delta^t \cdot \gamma_{i,t}(x_1, x_2, x_3, ...)$$

$$\le \sum_{i=0}^{\infty} S_i \sum_{t=i}^{\infty} \delta^t \cdot \gamma_{i,t}(X_z, 0, 0, ...)$$

$$= \Pi(X_z, 0, 0, ...)$$

We have established that without loss of optimality, the provider can increase inconvenience in one step. We can find an optimal solution by maximizing $\Pi(x, 0, 0, ...)$. Since $\gamma_{i,t}(x, 0, 0, ...) = p(x) \cdot p(0)^{t-i} \cdot \pi(x)$ for any $i \le t$,

$$\Pi(x, 0, 0, ...) = p(x) \cdot \pi(x) \sum_{i=0}^{\infty} S_i \sum_{t=i}^{\infty} \delta^t \cdot p(0)^{t-i},$$

so it suffices to maximize $p(x) \cdot \pi(x)$. ∎

Proof of Lemma 1: We first note that for a twice differentiable function g, $(\ln g(x))''$ has the same sign as

$$g''(x) \cdot g(x) - (g'(x))^2.$$

Let $\overline{F}(x) \equiv 1 - F(x)$. Then, $(\ln \overline{F}(t(x)))''$ has the same sign as

$$\overline{F}''(t(x)) \cdot \overline{F}(t(x)) - (\overline{F}'(t(x)))^2 + \overline{F}(t(x)) \cdot \overline{F}'(t(x)) \frac{t''(x)}{(t'(x))^2}.$$

Since $\overline{F}$ is decreasing, if $t''(x) < 0$ and $\overline{F}$ is log-convex, then $\overline{F}(t(x))$ is log-convex. We conclude that under the assumptions of this lemma, $p(x) \equiv \overline{F}(-v(-x))$ is log-convex.

∎

Proof of Lemma 2: Since $f(x) = F'(x)$, we have that $(\log(1 - F(x)))'' \geq 0$ if and only if

$$-f'(x) \cdot (1 - F(x)) \geq f(x)^2.$$

∎

Proof of Lemma 3: Straightforward calculations show that

$$\Pi(x, z+1) - \Pi(x, z) = (p(x)\pi((z+1)x) - p(0)\pi(zx))p(x)^z \sum_{t=z+1}^{\infty} \delta^t p(0)^{t-(z+1)}.$$

Thus, $\Pi(z+1, x) > \Pi(z, x)$ if

$$p(x)/p(0) > \pi(zx)/\pi((z+1) \cdot x),$$

and $\Pi(z+1, x) < \Pi(z, x)$ if

$$p(x)/p(0) < \pi(zx)/\pi((z+1) \cdot x).$$

Moreover, since π is log-concave, $\pi(zx)/\pi((z+1) \cdot x)$ is increasing in z (for a fixed x). We conclude that $\Pi(z, x)$ is unimodal in z for a fixed x: it is increasing for $z < z^*(x)$ and decreasing for $z > z^*(x)$. It is thus maximized at $z^*(x)$. ∎