Selling to Consumers with Endogenous Types*

Jan Boone[†] and Joel Shapiro[‡]

April 24, 2009

Abstract

For many goods consumers' preferences change over time. In this paper, we examine a monopolist's optimal pricing schedule when current consumption can affect a consumer's valuation in the future and valuations are unobservable. We assume that consumers are anonymous, *i.e.* the monopolist can't observe a consumer's past consumption history. For myopic consumers, the optimal consumption schedule is distorted upwards, involving substantial discounts for low valuation types. This pushes low types into higher valuations, from which rents can be extracted. For forward looking consumers, there is a further upward distortion of consumption; low valuation consumers now have a strong interest in consumption in order to increase their valuations. Firms will find it profitable to educate consumers and encourage forward looking behavior, if consumers don't suffer large negative effects from consumption.

Keywords: endogenous types, addictive goods, price discrimination
J.E.L. codes: D42, D82, L12

*We thank Heski Bar-Isaac, Simon Board, Pascal Courty, Eric Danan, Doh-Shin Jeon, Jens Josephson, Thijs Van Rens, Huseyin Yildirim and audiences at Barcelona Jocs, Duke/Northwestern/Texas I.O. Conference, Econometric Society NASM 2007 and NAWM 2008, EARIE 2007, EUI, Fordham, Groningen, IIOC, NYU-Stern, Queen's, UPF, and WZB for helpful comments. We also thank Gonçalo Pina and Milos Bozovic for excellent research assistance. Boone gratefully acknowledges financial support from the Dutch National Science Foundation (Grant number 453.07.003). Shapiro gratefully acknowledges financial support from the Spanish government under SEJ2006-09993, the Barcelona GSE Research Network, and the Generalitat de Catalunya.

[†]CentER, TILEC, Tilburg University and CEPR. Email: j.boone@uvt.nl

[‡]Departament d'Economia i Empresa, Universitat Pompeu Fabra and CEPR. E-mail: joel.shapiro@upf.edu

1. Introduction

For many goods, consumers' preferences change over time. These changes may be related to learning how much one needs of the good, where one tries lower qualities at first to decide whether to upgrade. The changes may also come directly from current consumption increasing one's desire for more, as in the case of addictive goods. In this paper we analyze monopoly pricing under asymmetric information when consumers' current consumption influences their future enjoyment of the good.

There are two critical features to our analysis:

- *Types are determined by past actions:* more consumption today gives one a larger probability of having a higher valuation for the good in the next period. This is an intensive margin that presents a trade-off for the monopolist - with heterogeneous consumers and asymmetric information about their valuations, downwards distortion of consumption usually solves the adverse selection problem, but may now hurt the firm's future profits.

- *Consumers are anonymous:* this implies that the monopolist can't make current prices contingent on the consumption history of a customer. One might think of the monopolist as engaging in repeated relationships that remain anonymous, perhaps because the firm distributes through different retailers or posts prices that are available to all consumers. The monopolist must then choose prices that balance getting low valuation consumers 'hooked' while extracting surplus from high valuation consumers.

Using these elements we find the monopolist's optimal steady-state pricing schedule. We look at both the case where consumers are myopic (they only consider today's utility), and where they are perfectly forward looking. When types are observable, the monopolist increases quality substantially (compared to the static first best[1]) at the bottom of the valuation distribution to get consumers 'addicted'. We find that in the case of myopic consumers, the adverse selection problem may still distort the quality schedule downwards, but not as much as in the static case. In order to entice low valuation customers to begin upgrading, the monopolist may offer substantial discounts, which we define as pricing below cost. In the case of forward looking consumers, we find that the adverse selection problem actually pushes the quality schedule upwards (compared to the static first best outcome and the myopic solution). This is driven by the fact that low valuation forward looking consumers have a strong preference to consume and realize type upgrades. In addition, we show that it is profitable for the monopolist to educate consumers and make them forward looking. Note, of course, that we assume that there is no harmful aspect to increased consumption for the majority of the paper. When we extend the model to add a cost of too much consumption (as in the case of harmful addiction), we find that consumption is lowered and that the monopolist's incentives to educate the consumer decrease with this cost.

Examples of products for which current consumption affects future consumption abound. In software, users often begin using a stripped down version of a program (offered at a substantial

[1] The static first best is defined here as marginal (direct) utility equal to marginal costs for each type.

discount) before moving on the full version.[2] Websites provide basic access (usually for free) with advertisements for those only browsing, but offer a chance to pay for premium content and reduced advertising.[3] Video games, particularly Massively Multiplayer Online Role Playing Games (MMORPG), offer users free trials and the ability to build their characters so as to enjoy the game more, and potentially gain access to new levels. Higher end car manufacturers often have a low cost entry level car to attract customers to the brand and persuade them to "trade up" in the future.[4] Similarly, people connecting to the internet often start with a slow connection and then upgrade over time as they start using the internet more and more intensely. And of course, purely addictive substances such as caffeine or exercise have these properties. Extending our model (as we do in section 4.2.2) to harmful addictive goods adds cigarettes, drugs, and many other substances to the list.

Naturally, we must add a caveat to our examples. Previous authors have noted price discrimination in these markets and provided different explanations. The majority of the explanations rely on taste heterogeneity. What we provide is another, complementary explanation and a general framework in which to think about pricing goods whose consumption affects preferences.

In the dynamic mechanism design literature, papers generally either address agents whose types remain the same over time or agents whose type is redrawn from a random distribution in each period. When types are constant over time, short term contracts are influenced by the "ratchet effect": agents' consumption becomes a signal of their type. Freixas, Guesnerie, and Tirole (1985), Laffont and Tirole (1988), and Skreta (2006) analyze this problem. We avoid the signaling problem by assuming that the firm can't make its mechanism dependent on previous consumption. This allows us to focus on the interaction between consumption and type and makes the analysis more tractable. Townsend (1982) models dynamic contracts where agents' types are generated by a random process. His work was extended by Green (1987), Thomas and Worrall (1990), and Atkeson and Lucas (1992). The recent literature on dynamic optimal taxation also assumes that individuals' types are randomly re-drawn each period. A review of this literature is provided by Kocherlakota (2005).

There are a couple of models that incorporate a learning by doing effect on types that is similar to our consumption effect. Baron and Besanko (1984) describe a two period model of regulation where the firm's choice of R&D in period 1 can affect its type in period 2. This differs from our analysis in two ways. First, the firms' choice of R&D doesn't interact with the firm's first period type; in our model it is important that current consumption interacts with the consumer's valuation of the good. Second, they analyze the case where the regulator commits to the two period mechanism (to avoid the ratchet effect) but allows the instruments in period 2 to depend on reports of type from both period 1 and period 2. Lewis and Yildirim (2002) analyze a dynamic model where there is renegotiation each period, types are redrawn randomly

[2] Intuit's program Quicken was offered as Basic Quicken ($20) and Quicken Deluxe ($60). Wolfram Research sells both Mathematica and a student version of it (which has a lesser computational speed and reduced functionality). For more examples, see Shapiro and Varian (1998).

[3] For example www.salon.com and www.wallstreetjournal.com are newspaper websites which offer different levels of access at different prices.

[4] Keller (2003) cites the case of BMW and its 3-series automobiles, as well as Mercedes and its A-Class vehicles.

each period, and the learning by doing reduction of costs is random and doesn't interact with the firm's type. Gärtner (2007) examines a two-period model of learning by doing where first period types are known and second period types are a function of first period production and an unobservable component that affects the learning curve.

Lastly, our model has some similarities to the literature on rational addiction initiated by Becker and Murphy (1988). In our model, current consumption increases future consumption (in expectation), so the goods we discuss have the property of addiction. For most of our paper, we look at "beneficial" addictive goods in the sense that utility is increasing in the stock[5] of previous consumption. In section 4.2.2, we examine how the results change with harmful addictions. Our focus, nevertheless is quite different from Becker and Murphy (1988). We focus on monopoly price discrimination, while they assume consumers are homogeneous and prices are competitive. There are several follow up papers to Becker and Murphy (1988) which look at pricing, but only in the case of homogeneous consumers. These include Becker, Grossman, and Murphy (1994), Fethke and Jagannathan (1996), Showalter (1999), and Driskill and McCafferty (2001).

The paper is organized as follows: Section 2 sets up the model and finds the steady state distribution of types. Section 3 looks at the optimal quality schedule when consumers are myopic. Section 4 examines the solution when consumers are forward looking. Section 5 concludes.

2. The Model

Consider a monopolist that sells a product to consumers with utility function $u(x, n) = nx - p$ where $x \geq 0$ is the quality of the good, $n \in [n_0, n_1]$ is the valuation of the consumer for the product and p is the price paid for the quality x.[6] The monopolist's cost function is represented by $c(x)$, which is assumed to be increasing and convex. We assume that the outside option of all consumers is equal to zero.

The good has addictive properties in the sense that the amount x consumed today affects the consumer's type n tomorrow. The higher today's consumption, the higher the consumer's expected type tomorrow. Hence, for the firm, the density (distribution) function $f(n)$ $(F(n))$ over types $[n_0, n_1]$ is endogenous. By changing the quality vector $x(.)$, the firm changes the distribution of types.

In particular, we assume that consumers live in continuous time τ and at each instant they have a poisson arrival rate $\delta d\tau$ of a taste shock that can reset their type. For each type n that isn't affected by the taste shock, there is a probability $\phi(x(n))$ that its type increases (a type upgrade) by $dn = \alpha(n)d\tau$ over the next instant. We assume that $\phi'(x) \geq 0$, *i.e.* consuming more quality leads to a higher probability of increasing one's type. This is the learning-by-consuming or addiction effect that we described previously. We also assume that low n types can experience bigger upgrades than high n types, *i.e.* $\alpha'(n) \leq 0$. Furthermore, low types cannot overtake

[5]We do not actually have a "stock" variable, but we loosely interpret one's type as summarizing past consumption.

[6]Alternatively, we could interpret x as the quantity purchased of the good (with n being an individual's valuation per unit purchased).

higher types that also experience an upgrade, *i.e.* $d(n + \alpha(n)d\tau)/dn = 1 + \alpha'(n)d\tau \geq 0$. Finally, we assume that there is a finite upper bound n_1 on the taste parameter. This implies that $\alpha(n_1) = 0$ and no type can end up above n_1.

In the case where a taste shock does occur, the consumer switches from his current type n to a newly drawn type from an exogenously given distribution $G(.)$ over $[n_0, n_1]$ and corresponding density function $g(.)$. In general, this exogenous distribution may have positive density only on a subset of types $[n_0, \varepsilon]$, where $\varepsilon \leq n_1$. Below we sometimes focus on the case where ε is close to n_0. In that case, the shock δ leads to a reduction in n for most types.

We assume that the price-quality schedule doesn't depend on a consumer's consumption history. In addition, we assume that the firm is sufficiently patient and maximizes with respect to the steady state distribution. These assumptions are made for tractability reasons. Given that types are not randomly drawn each period, we would like to avoid the well known "ratchet effect". The complexity of a dynamic optimization problem over types and time also disappears with this formulation. Nevertheless, we also have a compelling motivation for not allowing the mechanism to be time dependent. The inability to identify past consumers and their purchasing patterns is a large concern for retailers. Internet retailers have an easier time identifying consumers' purchases, but attempts to condition prices on consumption history have brought substantial controversy[7].

To derive the steady state distribution, we first consider how the distribution of types $F(.)$ varies over time. The following equation shows this time variation over a short instant $d\tau$. To derive this, it is convenient to temporarily write the distribution as a function of both type n and time τ. The amount of consumers below type n at time $\tau + d\tau$ is given by

$$F(n, \tau + d\tau) = \delta d\tau G(n) + (1 - \delta d\tau)\left(F(n, \tau) - \int_{n-\alpha(n')d\tau}^{n} \phi(x(\nu))f(\nu, \tau)d\nu\right)$$

There is a probability $\delta d\tau$ that a type n gets a taste shock and is redrawn from $G(.)$. Hence each instant there are $\delta d\tau G(n)$ 'new' types below n. Of the part that is not affected by this taste shock, only the group who had types larger than $n - \alpha(n')d\tau$ (where n' is defined by $n' + \alpha(n')d\tau = n$) find that a type upgrade $\alpha(n')d\tau$ is big enough to move beyond n. A fraction $\phi(x(\nu))$ of these types experience such an upgrade and move above n. The integral subtracts exactly this group. Hence at $\tau + d\tau$ there are two types of consumers who used to be (at time τ) below n and are now above it - the group who experiences a taste shock and draws a type larger than n from $G(.)$, and the types $\nu \in (n - \alpha(n')d\tau, n]$ who get a type upgrade. The equation above can be rewritten as:

$$\begin{aligned}\frac{F(n, \tau + d\tau) - F(n, \tau)}{d\tau} &= \delta(G(n) - F(n)) - \frac{1}{d\tau}\int_{n-\alpha(n')d\tau}^{n} \phi(x(\nu))f(\nu, \tau)d\nu + \\ &\quad \delta\int_{n-\alpha(n')d\tau}^{n} \phi(x(\nu))f(\nu, \tau)d\nu\end{aligned}$$

[7]The most famous example is Amazon.com's attempt at dynamic pricing. When consumers and consumer groups learned that the same good was being offered at different prices depending on consumption history at Amazon, complaints forced Amazon to rescind the policy and offer rebates to those who had previously been affected (see "Amazon backs off on price-testing efforts", by Deborah Kong, *USA Today*, September 29, 2000).

We take the limit $d\tau \to 0$ and find (note that the third term on the right hand side converges to zero as $d\tau \to 0$):

$$\frac{dF(n,\tau)}{d\tau} = \delta(G(n) - F(n)) - \phi(x(n))f(n)\alpha(n)$$

The steady state distribution is constant over time, *i.e.* $\frac{dF(n,\tau)}{d\tau} = 0$. Thus we find that for each n it must be the case that:

$$\delta(G(n) - F(n)) = \alpha(n)\phi(x(n))f(n) \tag{1}$$

This equation says that in the steady state the net inflow of people to types below n due to a taste shock equals the outflow of types due to the upgrading effect.

3. Myopic Consumers

3.1. The Static Benchmarks

In this section we provide two benchmarks. In both, we consider a simplified model, where *both* the firm and the consumer are myopic. We define myopic as only being concerned about current period (instant) returns. Therefore the firm takes the distribution of types $F(.)$ as given. That is, it does not understand the relation between consumption and types given by the distribution in equation (1). The myopic consumer not only doesn't take into account future utility, he doesn't link current consumption to future happiness.

The firm maximizes:

$$\max_{x(.),p(.)} \int_{n_0}^{n_1} [p(n) - c(x(n))]f(n)dn$$

Taking into account that $u(n) = nx(n) - p(n)$, the objective function of the firm becomes:

$$\max_{x(.),u(.)} \int_{n_0}^{n_1} [-u(n) + nx(n) - c(x(n))]f(n)dn$$

In the static first best (SFB) framework, the firm can observe individuals' types. The solution is straightforward; the firm sets every agent's utility equal to zero (making the participation constraint bind) and sets the quality such that $n - c'(x^{SFB}(n)) = 0$. The quality $x^{SFB}(n)$ is increasing in n, as are prices.

In the static second best (SSB) framework, the firm can't observe the types of consumers. Thus, the monopolist is subject to the incentive compatibility (IC) constraint that each type n chooses the right $(x(n), p(n))$ combination from the menu $(x(.), p(.))$ offered.[8] Denoting the solution in this case by SSB, standard analysis yields

$$n - c'(x^{SSB}(n)) = \frac{1 - F(n)}{f(n)} \tag{2}$$

provided that $x^{SSB}(n)$ is nondecreasing in n to guarantee that the solution is incentive compatible (see Fudenberg and Tirole (1991: 261)). If x^{SSB} is decreasing in n over some range, the

[8] Using the revelation principle, we can indeed focus on such direct revelation mechanisms.

solution needs to be "ironed out" using the procedure described in Fudenberg and Tirole (1991: 303-306). We do not explicitly characterize the solution in this case, but assume throughout that $dx^{SSB}(n)/dn \geq 0$.

Although the firm disregards how the schedule affects the distribution of types, we haven't assumed this effect away. In equation (2), it is clear that the distribution also affects the optimal schedule. In section 3.3, we provide an example that solves for the optimal schedule given the distribution effect.

3.2. First Best with a Forward Looking Firm

We now allow the firm to maximize long run profits and take into account the type upgrading effect. We assume that the instantaneous discount rate is $\rho \in \langle 0, 1 \rangle$ and that types are observable. When the firm contracts with a consumer of type n, its discounted profits are:

$$\begin{aligned} \pi(n) &= (p(n) - c(x(n)))d\tau + (1 - (\rho + \delta)d\tau)\{(1 - \phi(x(n)))\pi(n) + \phi(x(n))\pi(n + \alpha(n)d\tau)\} \\ &\quad + \delta d\tau \int_{n_0}^{n_1} \pi(m)g(m)dm \end{aligned}$$

The first term represents the current profits from a consumer of type n. The second term is the profits if the consumer doesn't have a taste shock - with probability $\phi(x(n))$ he has a taste upgrade to $n + \alpha(n)d\tau$ and otherwise he remains the same type. The third term is the expected value of profits if the consumer has a taste shock. We note that the probabilities $(1 - (\rho + \delta)d\tau)$ and $\delta d\tau$ come from linearizing $e^{-(\rho+\delta)d\tau}$ and $e^{-\rho d\tau} - e^{-(\rho+\delta)d\tau}$, respectively. Manipulating the expression and dividing by $d\tau$ yields:

$$(\rho + \delta)\pi(n) = p(n) - c(x(n)) + \phi(x(n))\alpha(n)\pi'(n) + \delta \int_{n_0}^{n_1} \pi(m)g(m)dm$$

where we have used the Taylor approximation $\pi(n + \alpha(n)d\tau) - \pi(n) = \alpha(n)d\tau\pi'(n)$ since $d\tau$ is small. Using this expression we can write the discounted total profits by integrating over types:

$$\begin{aligned} (\rho + \delta) \int_{n_0}^{n_1} \pi(n)f(n)dn &= \int_{n_0}^{n_1} \{p(n) - c(x(n)) + \phi(x(n))\alpha(n)\pi'(n) \\ &\quad + \delta \int_{n_0}^{n_1} \pi(m)g(m)dm\}f(n)dn \end{aligned}$$

We can simplify the expression by using equation (1) for the steady state distribution of types, substituting for $\phi(x(n))\alpha(n)$:

$$\begin{aligned} (\rho + \delta) \int_{n_0}^{n_1} \pi(n)f(n)dn &= \int_{n_0}^{n_1} \{p(n) - c(x(n))\}f(n) + \delta(G(n) - F(n))\pi'(n)dn \\ &\quad + \delta \int_{n_0}^{n_1} \pi(m)g(m)dm \end{aligned}$$

After integrating by parts, we can rewrite in a very straightforward way:

$$\int_{n_0}^{n_1} \pi(n)f(n)dn = \frac{1}{\rho} \int_{n_0}^{n_1} \{p(n) - c(x(n))\}f(n)dn$$

Therefore the discounted total profits is simply the discounted total revenues minus the discounted total costs. This shouldn't be very surprising, given that in the steady state the distribution of types remains the same at each instant. To eliminate prices, once again we substitute the utility function, making profits equal to:

$$\frac{1}{\rho}\int_{n_0}^{n_1}\{nx(n)-u(n)-c(x(n))\}f(n)dn$$

Since the firm knows the type of each agent, it can make the participation constraint bind, setting $u(n)=0$ for all n. This makes the maximization problem (taking into account the effect on the steady state distribution of types):

$$\max_{x(.),F(.),f(.)}\int_{n_0}^{n_1}(nx(n)-c(x(n)))f(n)+\lambda(n)(F'(n)-f(n))+\mu(n)(\alpha(n)\phi(x(n))f(n)-\delta(G(n)-F(n)))dn$$

This is an optimal control problem. The state variables are $F(n)$ and $x(n)$ and the control variable is $f(n)$. The costate variable for $F(n)$ is $\lambda(n)$, and $\mu(n)$ is the Lagrange multiplier associated with the steady state distribution of types constraint. The endpoints of $F(n)$ are defined: $F(n_0)=0$ and $F(n_1)=1$. The first order conditions (Euler equations) for x, F and f can be written as follows.

$$n-c'(x(n)) = -\mu(n)\phi'(x(n))\alpha(n) \quad (3)$$
$$\lambda'(n) = \mu(n)\delta \quad (4)$$
$$nx(n)-c(x(n))-\lambda(n)+\mu(n)\alpha(n)\phi(x(n)) = 0 \quad (5)$$

Instead of setting the marginal profit on type n, $n-c'(x(n))$, equal to zero as in the static first best, the firm now takes into account the effect that more consumption of quality has on type upgrading. The second equation gives the differential equation for $\lambda(n)$. Finally, equation (5) equalizes marginal costs and benefits of a small increase in $f(n)$.

The fact that $\alpha(n_1)=0$ yields that $x(n_1)$ coincides with the static first best solution $(n_1-c'(x(n_1))=0)$. In addition, the condition $n_1x(n_1)-c(x(n_1))-\lambda(n_1)=0$ holds.

We can prove that the distortion of $x(n)$ is an upwards distortion from the static first best solution. The main step is to prove that $\mu(n)\geq 0$ or equivalently $nx(n)-c(x(n))-\lambda(n)\leq 0$. We denote the solution as $x^{DFB}(n)$, where DFB signifies "dynamic first best".

Proposition 1 *For all* $n\in[n_0,n_1\rangle$, $x^{DFB}(n)>x^{SFB}(n)$.

Proof We begin by proving that $nx(n)-c(x(n))-\lambda(n)\leq 0$ for all n. Suppose not, *i.e.* $nx(n)-c(x(n))-\lambda(n)>0$ for some n. Using equation (5), we see from equation (4), $\lambda'(n)<0$ and from equation (3), $n-c'(x(n))>0$. Therefore the derivative with respect to n of $nx(n)-c(x(n))-\lambda(n)$, $x(n)+(n-c'(x))\frac{dx}{dn}-\lambda'$, is positive. Since, $n_1x(n_1)-c(x(n_1))-\lambda(n_1)=0$, we find a contradiction (it can't be that $nx(n)-c(x(n))-\lambda(n)$ is positive and increasing for all n).

Now suppose $nx(n)-c(x(n))-\lambda(n)=0$ for any $n\in[n_0,n_1\rangle$. This implies that $\mu(n)=0$. Taking the derivative of $nx(n)-c(x(n))-\lambda(n)=0$ yields the equation $n-c'(x(n))=\frac{-x(n)}{x'(n)}$, using

$\mu(n) = 0$ and equation (4). However, since $\mu(n) = 0$, equation (3) gives us $n - c'(x(n)) = 0$. This gives us a contradiction.

Therefore, for all $n \in [n_0, n_1\rangle$, $nx(n) - c(x(n)) - \lambda(n) < 0$ and $\mu(n) > 0$. Using this fact and equation (3) proves the proposition. *Q.E.D.*

This demonstrates that the dynamic first best sells higher qualities to all consumers (except the highest valuation consumer) than the static first best. This is intuitive since the firm now takes into account the fact that raising $x(n)$ makes it more likely that type $n < n_1$ upgrades to a higher (and more profitable) type. In addition, the quality schedule is strictly greater than the static second best schedule for all n except for n_1.

If, in this case, a quality level $\bar{x}$ was reached where any additional quality changes would have no effect in changing the probability of upgrading (*i.e.* $\phi'(x) = 0$ for $x > \bar{x}$ and $\bar{x} < x^{DFB}(n_1)$), the upward distortion of quality would disappear. That is, if $x(\hat{n}) = \bar{x}$, then for all $n > \hat{n}$, qualities would be equal to those of the static first best.

3.3. Second Best with a Forward Looking Firm

We now assume that the forward looking monopolist doesn't observe the valuations of the myopic consumers. The monopolist's task is to encourage consumption and make consumers realize these type enhancements in order to extract rents from them while taking into account incentives for consumers to misstate their type. Incentive Compatibility is given by $u'(n) = x(n)$. Setting $u(n_0) = 0$, as there is no reason to leave surplus to the lowest type, allows us to solve $u(n) = \int_{n_0}^{n} x(t)dt$. The price can then be solved from $u(n) = nx(n) - p(n)$. Using integration by parts, the optimal control problem for the firm can be written as

$$\max_{x(.),F(.),f(.)} \int_{n_0}^{n_1} [nx(n) - c(x(n))]f(n) - (1 - F(n))x(n) +$$
$$\lambda(n)(F'(n) - f(n)) + \mu(n)(\alpha(n)\phi(x(n))f(n) - \delta(G(n) - F(n))dn$$

As above, the endpoints of $F(n)$ are defined: $F(n_0) = 0$ and $F(n_1) = 1$. The first order conditions (Euler equations) for x, F and f can be written as follows.

$$n - c'(x(n)) = \frac{1 - F(n)}{f(n)} - \mu(n)\alpha(n)\phi'(x(n)) \quad (6)$$
$$\lambda'(n) = x(n) + \delta\mu(n) \quad (7)$$
$$nx(n) + \mu(n)\alpha(n)\phi(x(n)) = c(x(n)) + \lambda(n) \quad (8)$$

The first equation equates the margin $n - c'(x(n))$ to the informational rent and the type upgrading margin (which we saw in the dynamic first best). The second equation gives the differential equation for $\lambda(n)$. Equation (8) equalizes marginal costs and benefits of a small increase in $f(n)$.

The shape of the optimal quality schedule can be compared easily to our benchmarks. Consider $n = n_1$. Since $\alpha(n_1) = 0$, it must be that $n_1 - c'(x(n_1)) = 0$. This proves that there is a "no distortion at the top" result, where the quality consumed by the highest type here is equivalent to that consumed by the highest type in the dynamic first best (as well as the static first best and the static second best). We denote the solution as $x^M(n)$, where M stands for myopic. As with the SSB case above, for $x^M(n)$ to be incentive compatible, we need $dx^M(n)/dn \geq 0$. If the solution generated by the equations above is decreasing in n over some range, it needs to be "ironed out" to get x^M. We assume that $dx^M(n)/dn \geq 0$ holds for all n and below derive conditions under which x^M is strictly increasing in n. We now prove that the myopic quality level is weakly larger than the static second best for all n, and strictly larger for a non-empty interval.

Proposition 2 *Assume that there exists $\tilde{n} < n_1$ such that $G(\tilde{n}) = 1$. Then for all $n \geq \tilde{n}$ the quality schedule $x^M(n) \geq x^{SSB}(n)$. Furthermore, there is a non-empty interval where $x^M(n) > x^{SSB}(n)$.*

Proof This follows from equation (6) and the following inequality:

$$nx(n) - c(x(n)) - \lambda(n) \leq 0 \tag{9}$$

with strict inequality for some non-empty interval, because this implies that $\mu(n) \geq 0$ with strict inequality for some non-empty interval. We also note that equation (8) implies that $n_1 x(n_1) - c(x(n_1)) - \lambda(n_1) = 0$.

To show that (9) holds, we first prove that the weak inequality holds by contradiction. Suppose not, *i.e.* suppose there exists n' such that $n'x(n') - c(x(n')) - \lambda(n') > 0$. Isolate $\mu(n') < 0$ in equation (8) and substitute it into equation (7). This implies that $\lambda'(n') < x(n')$. Hence we find

$$\frac{d(n'x(n') - c(x(n')) - \lambda(n'))}{dn'} > (n' - c'(x(n')))x'(n') \geq 0$$

where the last inequality follows from $n' - c'(x(n')) \geq 0$ (see equation (6)) and $x'(n') \geq 0$. However, $nx(n) - c(x(n)) - \lambda(n) > 0$ and increasing in n contradicts $n_1 x(n_1) - c(x(n_1)) - \lambda(n_1) = 0$.

Next, we prove that the strict inequality (in (9)) holds for some non-empty interval. If it didn't, then $nx(n) - c(x(n)) - \lambda(n) = 0$ for all n. This implies that $\mu(n) = 0$ for $n \in \langle n_0, n_1 \rangle$. Taking the derivative of $nx(n) - c(x(n)) - \lambda(n) = 0$ yields the equation $(n - c'(x(n)))x'(n) = 0$, using $\mu(n) = 0$ and equation (7). But $\mu(n) = 0$ also implies $c'(x(n_0)) < n_0$. Together with $x'(n) = 0$ and $c'(x(n_1)) = n_1 > n_0$ this yields a contradiction. *Q.E.D.*

The monopolist pushes up the qualities consumed relative to the static second best in order to take advantage of the boost in demand created for its product from high types. This increase in quality doesn't affect the whole schedule. At the top, the consumer is kept at the static efficient point. There is little reason to distort the consumption upwards at the top since this type is not able to benefit from the upgrading effect.

This intuition allows us to explore the schedule further. Imagine that for large enough consumption $x > \bar{x}$, there was no effect in changing the probability of upgrading, *i.e.* $\phi'(x) = 0$. Also suppose that $x^M(n_0) < \bar{x} < x^M(n_1)$, meaning that for types above some cutoff $\hat{n}$, there is no longer a marginal upgrading effect. Then for all $n > \hat{n}$, $x^M(n) = x^{SSB}(n)$ (this can be seen from equation (6)). That is, when there is no marginal upgrading effect, the upwards distortion on consumption is not optimal for the monopolist.

The advantage of the assumption in the above proposition on the cutoff $\tilde{n}$ (where for all $n \geq \tilde{n}$ we have $G(n) = 1$) is that it yields a very useful simplification. In particular, for $n \geq \tilde{n}$, the hazard rate $\frac{1-F(n)}{f(n)}$ equals $\frac{\alpha(n)\phi(x(n))}{\delta}$ and we can directly characterize the quality schedule. We use this property once again in the following lemma to show that the quality schedule is strictly increasing.

Lemma 1 *Assume that there exists $\tilde{n} < n_1$ such that $G(\tilde{n}) = 1$. Then $\phi''(.) \leq 0$ implies that $dx^M(n)/dn > 0$ for all $n \geq \tilde{n}$.*

Proof Substituting for $\frac{1-F(n)}{f(n)}$ and $\mu(n)$ in equation (6) and using the implicit function theorem we find that

$$\frac{dx^M(n)}{dn} = \frac{1 - \frac{\alpha'\phi}{\delta} + \frac{\phi'}{\phi}(\lambda' - x)}{c'' + \frac{\alpha\phi'}{\delta} - \frac{\phi''\phi - \phi'^2}{\phi^2}(\lambda - nx + c(x)) + \frac{\phi'}{\phi}(n - c'(x))}$$

We prove that this expression is positive. First, note that the numerator is positive as $\alpha'(.) \leq 0, \phi'(.) \geq 0$ and $\lambda'(n) - x(n) = \delta\mu(n)$ which is non-negative by equations (8) and (9). Next, turn to the denominator:

$$c'' + \frac{\alpha\phi'}{\delta} - \frac{\phi''\phi - \phi'^2}{\phi^2}\mu(n)\alpha(n)\phi(x(n)) + \frac{\phi'}{\phi}(n - c'(x))$$

Re-writing:

$$c'' + \frac{\alpha\phi'}{\delta} - \phi''\mu(n)\alpha(n) + \frac{\phi'}{\phi}\mu(n)\alpha(n)\phi'(x(n)) + \frac{\phi'}{\phi}(n - c'(x))$$

We can combine terms to get:

$$c'' + \frac{\alpha\phi'}{\delta} - \phi''\mu(n)\alpha(n) + \frac{\phi'}{\phi}(\frac{\alpha(n)\phi(x(n))}{\delta})$$

Which is then:

$$c'' + \frac{2\alpha\phi'}{\delta} - \phi''\mu(n)\alpha(n) > 0$$

which is positive given that $c'' \geq 0$ and $\phi'' \leq 0$. *Q.E.D.*

Finally, consider the following example to compare the myopic outcome with the schedules derived before.

Example 1 *Consider the case where $[n_0, n_1] = [10, 11]$ and $G(n) = n - 10$. Further assume that $\alpha(n) = 0.1(11 - n)$, $\phi(x) = x$, $\delta = 1$ and $c(x) = \frac{1}{2}x^2$. The static second best solution is given by the differential equation for $F(n)$ (where $F(n_0) = 0$) and the solution for $x(n)$:*

$$\begin{aligned} f(n) &= \frac{n - 10 - F(n)}{0.1(11-n)x(n)} \\ x(n) &= n\frac{F(n) - n + 10}{F(n) - n + 10 + 0.1(11-n)(F-1)} \end{aligned}$$

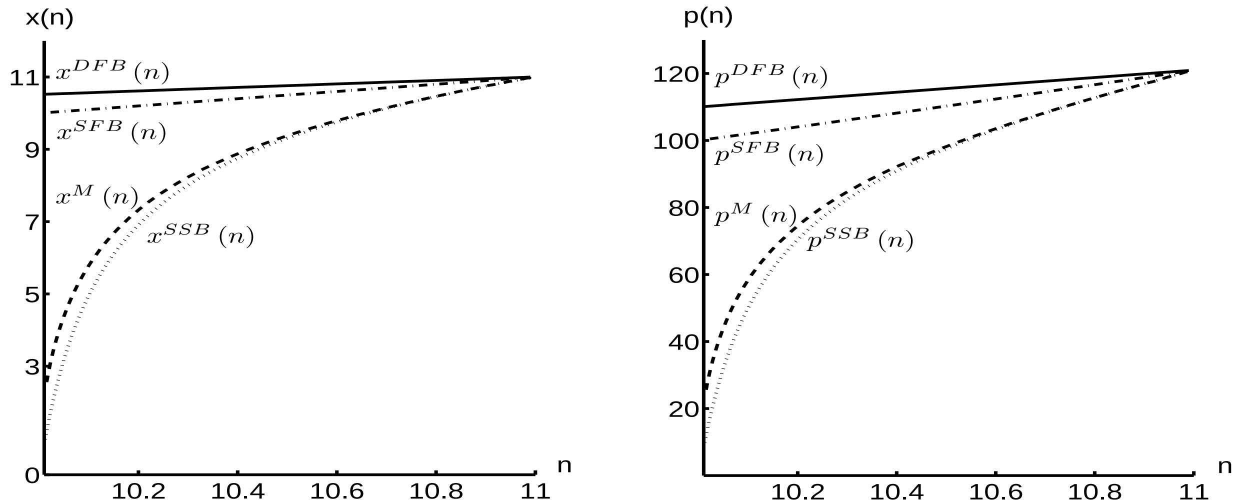

Figure 1: *Quality $x(n)$ (left) and price $p(n)$ (right) as a function of the consumer's type n.*

In figure 1 the dotted curve gives the solution to x^{SSB}. We also plot the static first best solution x^{SFB}, which is strictly larger than x^{SSB}, except at the top of the valuation distribution.

For $x^{DFB}(n)$, we use the necessary conditions:

$$\mu(n) = \frac{x(n)-n}{\alpha(n)}$$
$$\lambda(n) = \frac{1}{2}(x(n))^2$$

Thus we find

$$\lambda'(n) = x(n)\,x'(n)$$

Combining this with equation (9) we find the following differential equations for $x(n)$ and $F(n)$:

$$x'(n) = \frac{1}{\alpha(n)}\left(1 - \frac{n}{x(n)}\right)$$
$$f(n) = \frac{n-10-F(n)}{\alpha(n)x(n)}$$

With boundary condition $x(n_1) = n_1$ and initial condition $F(n_0) = 0$. The solution of this differential equation, $x^{DFB}(n)$, is the solid line in figure 1.

For $x^M(n)$, we use the necessary conditions (substituting for $\mu(n)$):

$$f(n) = \frac{n-10-F(n)}{0.1(11-n)x(n)}$$
$$0 = \frac{1-F(n)}{f(n)} + (x(n)-n) - (\lambda'(n) - x(n))\,0.1\,(11-n)$$
$$\lambda'(n) = \frac{1}{2}\frac{x(n)}{0.1(11-n)} - \frac{n}{0.1(11-n)} + \frac{\lambda(n)}{0.1(11-n)x(n)} - x(n)$$

together with the boundary equations $x(n_1) = n_1$, $\lambda(n_1) = \frac{1}{2}(n_1)^2$ and initial condition $F(n_0) = 0$. The solution to this system of differential equations, $x^M(n)$, is given by the dashed line in figure 1. The figure illustrates that $x^M(n) \geq x^{SSB}(n)$, with a strict inequality for an interval.

The prices are easy to solve for. Since the individual rationality constraint binds for the first best cases, $p^{SFB}(n) = nx^{SFB}(n)$ and $p^{DFB}(n) = nx^{DFB}(n)$. In the second best cases, we use the incentive compatibility condition $u(n) = \int_{n_0}^{n} x(t)dt$ to find $p^{SSB}(n) = nx^{SSB}(n) - \int_{n_0}^{n} x^{SSB}(t)dt$ and $p^M(n) = nx^M(n) - \int_{n_0}^{n} x^M(t)dt$. As shown in figure 1, the ordering of the prices is the same as the ordering of the quantities.

3.4. Discounts and prices below cost

A natural question arises when the monopolist increases quality for low types to encourage upgrading: what do prices look like? Do we observe the phenomena of discounting? Discounts (or introductory offers) often arise in the literature on experience goods[9]; however the main explanation given for the use of discounts is the presence of quality uncertainty by consumers. Low prices encourage experimentation and can boost future demand for the firm. Our explanation, that low prices promote type upgrading, encompasses this story and allows for addiction effects as well.

The brief answer to our question is that prices can be very low in our model, and in what follows we describe an example where prices can actually be below cost for low types.[10] The monopolist sells at a loss in order to create higher types to extract rents from.

Assume the function $\phi(x)$ takes the form:

$$\phi(x) = \begin{cases} 0 \text{ if } x < \bar{x} \\ \phi \text{ if } x \geq \bar{x} \end{cases}$$

where $\bar{x}$ is the quality level at which upgrading begins. Hence we assume that a consumer who consumes low quality doesn't get a taste for what the high end good could yield. Therefore, such consumption of low quality does not generate an upgrade. An internet user with a poor connection, for example, may never experience the convenience of downloading music and will not be tempted to upgrade her connection.

Furthermore, let the exogenous distribution $G(n)$ be a uniform distribution on $[n_0, n_{1g}]$, *i.e.* $G(n) = \frac{n-n_0}{n_{1g}-n_0}$, and the cost function be quadratic, $c(x) = \frac{1}{2}x^2$. The following assumption reduces the number of cases that we need to consider below.

$$n_{1g}\bar{x} - c(\bar{x}) = 0$$

This equation implies $\bar{x} = 2n_{1g}$.

This equation says that if the firm decides to sell $\bar{x}$ to types $n < n_{1g}$, it sells at a price below costs. Since the maximum one could charge type n_{1g} would be $n_{1g}\bar{x}$ (due to the participation constraint), any price below that would lose the firm money. The firm breaks even by extracting

[9]For example, see Shapiro (1983), Schlee (2001), and Gabszewicz, Pepall, and Thisse (1992). Price discrimination in these models is intertemporal. Our model has both static (*i.e.* within period) and intertemporal price discrimination.

[10]Although figure 1 displays that prices are actually lower in the static second best case than the myopic case, it is straightforward to prove that prices can't be below cost in the static second best case.

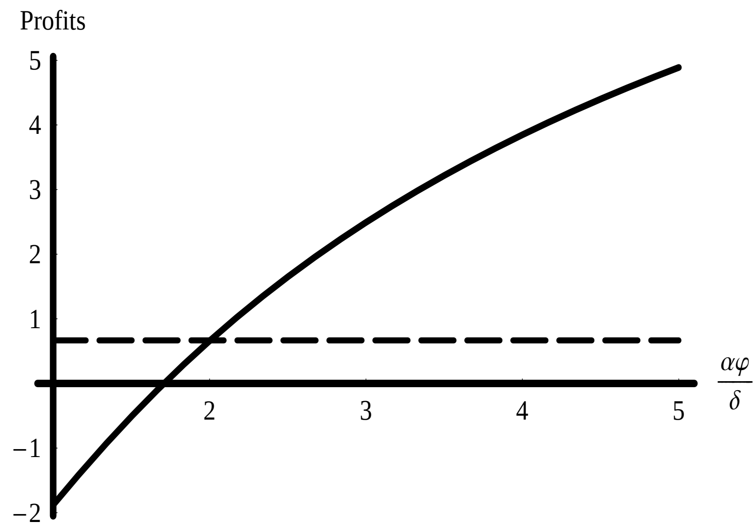

Figure 2: *Profits as a function of* $\frac{\alpha\phi}{\delta}$ *for the case without (dash) and with (solid) upgrading.*

all of type n_{1g}'s surplus, and therefore would lose money if it encouraged lower types to consume $\bar{x}$ and upgrade. This leaves the firm with two options: sell qualities below $\bar{x}$ to all consumers and have no upgrades, or to sell high quality goods and encourage upgrades. The second option, which by construction involves prices below costs, will be more profitable for certain parameters.

First, we examine the case where the firm sells qualities below $\bar{x}$ and no upgrading occurs. The optimal quality schedule is given by:

$$n - c'(x(n)) = \frac{1 - G(n)}{g(n)}$$

Since $x(n) < \bar{x}$ for all $n \in [n_0, n_{1g})$, there will not be any upgrading. With the functions that we chose and assuming in addition that $2n_0 \geq n_{1g}$, we find that:

$$\begin{aligned} x(n) &= 2n - n_{1g} \\ p(n) &= nx(n) - \int_{n_0}^{n} x(t)dt \\ \text{profit} &= \int_{n_0}^{n_{1g}} [p(n) - c(x(n))]g(n)dn = (4n_0^2 - 2n_0 n_{1g} + n_{1g}^2)/6 \end{aligned}$$

Second, we look at a quality schedule where upgrading occurs. The schedule we choose may not be the optimum, but since we prove that it can yield higher profits than the case without upgrading, the optimum (for those parameters) must involve upgrading. Moreover, since any case with upgrading involves pricing below cost, deep discounts will be a part of the optimal solution (for those parameters).

Suppose all types in the interval $[n_0, n_{1g}]$ are sold $\bar{x}$. It follows that $p(n) = n_0\bar{x}$ for these types, otherwise type n_0 would not buy. Since $n_0\bar{x} - c(\bar{x}) < 0$ the firm makes a loss on these consumers. And since $x(n)$ cannot fall (by incentive compatibility) and there is no reason to sell qualities in excess of $\bar{x}$, we find $x(n) = \bar{x}$ for all $n \in [n_0, \bar{n}]$ where $\bar{n}$ is defined by

$$\bar{n} - c'(\bar{x}) = \frac{1 - F(\bar{n})}{f(\bar{n})}$$

Thus, the quality schedule for types in the interval $[\bar{n}, n_1]$ is given by $n - c'(x) = \frac{1-F(n)}{f(n)}$. We assume that $\alpha(n) = \alpha \cdot (n_1 - n)$ and show how $F(n)$ and $\bar{n}$ are determined in the appendix. To types $n > \bar{n}$ the firm sells

$$x(n) = (1 + \frac{\alpha\phi}{\delta})n - \frac{\alpha\phi}{\delta}n_1$$

From this we can calculate the profits for $n > \bar{n}$ in the standard way.

Now we can analyze whether it is optimal for the firm to sell $\bar{x}$ to the low types at a loss such that the upgrading can take effect. We consider the parameters $n_0 = 1, n_{1g} = 2$, and $n_1 = 10$. Profits in the no-upgrading solution (when $x(n) < \bar{x}$ for all n) are then equal to 2/3. Selling $\bar{x}$ to low types leads to profits as depicted in figure 2. Clearly for $\alpha\phi/\delta$ high enough, it is optimal for the firm to sell to low types at a loss. This is intuitive: the stronger the upgrading effect (because $\alpha\phi/\delta$ is high) the more profitable the strategy that induces upgrading.

4. Results on Forward Looking Consumers

4.1. Continuous Types

In this section, we assume that consumers understand perfectly the type upgrading process and take it into account when making consumption decisions. To characterize incentive compatibility in this case, we first derive the expected discounted value of a type n consumer. Denoting this value by $V(n)$, we can write

$$\begin{aligned} V(n) &= (nx(n) - p(n))d\tau + (1 - (\rho + \delta)d\tau)((1 - \phi(x(n)))V(n) + \\ &\quad \phi(x(n))V(n + \alpha(n)d\tau)) + \delta d\tau \int_{n_0}^{n_1} V(m)g(m)dm \end{aligned}$$

Hence the expected discounted value of being of type n consists of three parts. First, the consumer receives the instant utility $nx(n) - p(n)$ over the short 'period' $d\tau$. Second, with some probability the consumer experiences no taste shock and gets a type upgrade of $\alpha(n)d\tau$ with probability $\phi(x(n))$, or continues with payoff $V(n)$ with probability $1 - \phi(x(n))$. Third, if the consumer experiences the taste shock, the payoff is the expectation of the continuation value $V(m)$, where m is distributed according to the exogenously given density function $g(m)$. Using the standard techniques of deriving a Bellman equation we get the following equation for $V(n)$:

$$(\rho + \delta)V(n) = nx(n) - p(n) + \alpha(n)\phi(x(n))V'(n) + \delta \int_{n_0}^{n_1} V(m)g(m)dm \qquad (10)$$

Rewriting to isolate the price $p(n)$:

$$p(n) = nx(n) - (\rho + \delta)V(n) + \alpha(n)\phi(x(n))V'(n) + \delta \int_{n_0}^{n_1} V(m)g(m)dm \qquad (11)$$

We will use this to eliminate prices from the objective function of the firm. Integrating over types (and using equation (1) and integration by parts) yields:

$$\int_{n_0}^{n_1} p(n)f(n)dn = \int_{n_0}^{n_1} (nx(n) - \rho V(n))f(n)dn$$

Expected discounted profits for the firm can then be written as:

$$\int_{n_0}^{n_1} \pi(n)f(n)dn = \frac{1}{\rho}\int_{n_0}^{n_1} (nx(n) - c(x(n)) - \rho V(n))f(n)dn \tag{12}$$

In the first best solution, the firm maximizes profits subject to the equation representing the steady state distribution of types and the participation constraint. Since types are observable, the firm can set $V(n) = 0$ for all n. This then makes the problem equivalent to the dynamic first best problem when agents are myopic. The solution therefore is the dynamic first best quality schedule, $x^{DFB}(n)$. We formalize this in a proposition:

Proposition 3 *The first best solution with forward looking consumers is equivalent to the first best solution with myopic consumers, $x^{DFB}(n)$.*

Clearly, the key to this proof is that the outside option is the same (and equal to zero) for both myopic and forward looking consumers.

We now examine the second best solution.

To derive incentive compatibility we take one step back and write:

$$(\rho + \delta)V(n) = \max_{\tilde{n}}\{nx(\tilde{n}) - p(\tilde{n}) + \alpha(n)\phi(x(\tilde{n}))V'(n) + \delta\int_{n_0}^{n_1} V(m)g(m)dm\} \tag{13}$$

Using the envelope theorem we find the incentive compatibility condition:

$$(\rho + \delta)V'(n) = x(n) + \alpha'(n)\phi(x(n))V'(n) + \alpha(n)\phi(x(n))V''(n) \tag{14}$$

It is straightforward to see that the solution to the first best problem doesn't hold here. Suppose $V(n) = 0$ for the whole interval $[n_0, n_1]$. This violates the incentive compatibility constraint since it produces the equation $0 = x(n)$ for all n. Since the monopolist could make positive profits by selling small amounts of the good, it can't be that $x(n) = 0$ everywhere.

The following result gives us insight into the value function.

Proposition 4 *For functions $V(.)$ that are twice differentiable we find that $V'(n) \geq 0$.*

The proof is in the appendix.

Substituting the fact that $V(n) = V(n_0) + \int_{n_0}^{n} V'(n)dn$ into the firm's objective function given by equation (12) and integrating by parts gives the firm's maximization problem:

$$\max_{x(.),F(.),f(.),\nu(.),\omega(.)} -V(n_0) + \int_{n_0}^{n_1} [nx(n) - c(x(n))]f(n) - \rho(1 - F(n))\nu(n) +$$
$$\lambda(n)(F'(n) - f(n)) + \mu(n)(\alpha(n)\phi(x(n))f(n) - \delta(G(n) - F(n)) + \psi(n)(\nu'(n) - \omega(n))$$
$$+\xi(n)[(\rho + \delta - \alpha'(n)\phi(x(n))\nu(n) - x(n) - \alpha(n)\phi(x(n))\omega(n)]dn$$

where $\nu(n) = V'(n), \omega(n) = \nu'(n) = V''(n)$, $\psi(.)$ is the costate variable associated with $\nu(n)$, and $\xi(.)$ is a Lagrange multiplier. The firm is constrained by the steady state distribution of types equation and incentive compatibility. It is clear from the problem itself that the value function for the lowest type will be set equal to zero (the participation constraint binds and $V(n_0) = 0$). In this case, there are five optimality conditions. We list them all in the appendix, and only focus on the one with respect to $x(n)$ here:

$$(n - c'(x(n)))f(n) = -\mu(n)\alpha(n)\phi'(x(n))f(n) + \xi(n)(1 + \alpha'(n)\phi'(x(n))\nu(n) + \alpha(n)\phi'(x(n))\omega(n)) \tag{15}$$

The right hand side, which is the wedge between this solution and the static first best one, consists of two terms. The first is the type upgrading effect, which takes the same form as in the myopic case. The second is the distortion due to asymmetric information and the need for incentive compatibility. We now examine this second term. In the following proposition we prove that the term without the multiplier,

$$\frac{d}{dn}\left(-\frac{dV/dx}{dV/dp}\right) = \frac{d(n + \alpha(n)\phi'(x)V'(n))}{dn} = 1 + \alpha'(n)\phi'(x(n))\nu(n) + \alpha(n)\phi'(x(n))\omega(n)$$

is positive.

Proposition 5 *Assume that $\phi''(.) \leq 0$. Then*

$$1 + \alpha'(n)\phi'(x(n))V'(n) + \alpha(n)\phi'(x(n))V''(n) \geq 0. \tag{16}$$

Proof

Substituting for $V''(n)$ from equation (14), we write this inequality as

$$1 - \frac{\phi'(x)x}{\phi(x)} + V'(n)(\rho + \delta)\frac{\phi'(x)}{\phi(x)} \geq 0$$

This inequality holds, since $V'(n) \geq 0$ and $\phi(x) \geq \phi'(x)x$. To see the last inequality we use the following Taylor series

$$\phi(0) = \phi(x) + \phi'(x)(0 - x) + \frac{1}{2}\phi''(\zeta)(0 - x)^2$$

for some $\zeta \in [0, x]$. This can be rewritten as

$$\phi(x) = \phi(0) + \phi'(x)x - \phi''(\zeta)x^2 \geq \phi'(x)x$$

Q.E.D.

In the appendix we prove that the multiplier $\xi(n)$ must be non-positive (see corollary 2). Looking at equation (15), using Proposition 5 and $\xi(n) \leq 0$, we can then conclude that the contribution of incentive compatibility is to increase the quality schedule, rather than reduce it (as it does in the myopic case and the static second best).

This reversal is explained by the difference between myopic and forward looking consumers. To dissuade high types from masquerading as low types in the case with myopic consumers (and in the static case as well), the quality schedule is generally distorted downwards, since high types have a strong preference for quality over low types. Here however, low types have a strong preference for quality as well, since they benefit both from the quality now and the type upgrades in the future.

Nevertheless, the "no distortion at the top" result present in the previous models continues to hold here. In equation (15), $\alpha(n_1) = 0$ eliminates the first term of the right hand side. We also prove in the appendix that $\xi(n_1) = 0$. This eliminates the second term on the right hand side of equation (15), showing that quality is set such that the marginal cost equals the valuation for the top type, without any distortion.

We now prove an intriguing result: the firm is better off if the consumer becomes forward looking instead of myopic. In reality, we see firms recognizing this effect: firms selling running shoes have web sites with information about running, where one can keep a running diary and have access to other training aids. Car makers encourage test drives of their luxury models and hype their features. When the consumer can realize the gains of upgrading, the firm has less need to push higher quality onto low types at low prices and can make a larger profit.[11]

Proposition 6 *Total profits are larger when consumers are forward looking.*

The proof is in the appendix. Clearly this result depends on the assumption that consumers can only benefit from type upgrades. Since their preferences are more aligned with those of the firm, the firm is better off.[12] In the myopic case, the firm has to offer incentives to get consumers to purchase more and upgrade, whereas forward looking consumers do not need to be convinced of the future benefits.

This raises two important questions. First, how does the myopic schedule compare to the forward looking schedule given this difference in the cost of the firm in providing incentives to upgrade? Second, does this result change when there are negative effects of addiction? We address these questions, using a two type version of the model in the next subsection. Using two types allows us to get closed form results that permit interesting comparative statics.

4.2. Comparative Statics in the two type model

Consider a model with two types $\{n_l, n_h\}$ where $n_l < n_h$. The dynamics are assumed to be as follows. With probability $\delta d\tau$, a taste shock hits at each moment in time and converts a type n_h into a type n_l. With probability $\phi(x_l)d\tau$, a type n_l receives a type upgrade and becomes a type n_h. We assume that $\phi(0), \phi'(.) \geq 0$ and $\phi''(.) \leq 0$.

[11] This might suggest that consumers have the incentive to pretend to be myopic. In a previous version of the paper, we discuss an example which shows that there is a tradeoff for consumers (*i.e.* their incentives are not unambiguous). Two effects are present. Since the lowest type has larger rents (to be extracted) when forward looking, it is advantageous to pretend to be myopic. On the other hand, the interests of the monopolist and the forward looking agent may be aligned more than that of the monopolist and the myopic agent in the sense that both prefer upgrades. Hence, pretending to be myopic may lead to less upgrades and a disappearance of surplus (for both the monopolist and the agent).

[12] In section 4.2.2, we demonstrate that this result may not be true when an addiction is harmful.

We write the steady state equation by first defining $\theta(\tau)$ as the proportion of type n_l consumers at time τ. This implies:

$$\theta(\tau + d\tau) = \delta d\tau(1 - \theta(\tau)) + (1 - \phi(x_l)d\tau)\theta(\tau)$$

Re-arranging and letting $d\tau \to 0$ gives us:

$$\theta = \frac{\delta}{\delta + \phi(x_l)} \tag{17}$$

All of the qualitative results for the previous model have analogues in this model.[13]

4.2.1. Comparing Myopic and Forward Looking Consumers

In order to facilitate the comparison between myopic consumers and forward looking consumers, we define $\Delta \in [0, 1]$ as the degree to which consumers understand the upgrade $\phi(x)$. Setting $\Delta = 0$ corresponds to a situation where consumers are myopic in the sense that they do not understand that they can upgrade type (but still value the future). Consumers who are fully forward looking understand that they will upgrade with probability $\phi(x)$; this corresponds to a value $\Delta = 1$.

The value function for a low type consumer can be written as:

$$V(n_l) = (n_l x_l - p_l)d\tau + (1 - \rho d\tau)((1 - \Delta\phi(x_l)d\tau)V(n_l) + \Delta\phi(x_l)d\tau V(n_h))$$

Taking $d\tau \to 0$ and rearranging:

$$\rho V(n_l) = n_l x_l - p_l + \Delta\phi(x_l)(V(n_h) - V(n_l))$$

Similarly, for the high type, the value function takes the form:

$$\rho V(n_h) = n_h x_h - p_h + \delta(V(n_l) - V(n_h))$$

Solving for $V(n_l), V(n_h)$, we get:

$$\begin{aligned} V(n_l) &= \frac{1}{\rho + \Delta\phi(x_l) - \frac{\delta\Delta\phi(x_l)}{\rho+\delta}}\left(n_l x_l - p_l + \frac{\Delta\phi(x_l)}{\rho + \delta}(n_h x_h - p_h)\right) \\ V(n_h) &= \frac{1}{\rho + \delta - \frac{\delta\Delta\phi(x_l)}{\rho+\Delta\phi(x_l)}}\left(n_h x_h - p_h + \frac{\delta}{\rho + \Delta\phi(x_l)}(n_l x_l - p_l)\right) \end{aligned}$$

[13] Specifically, the results are (i) $x_l^{DFB} > x_l^{SFB}$, (ii) the first best is the same in the myopic and forward looking cases, (iii) $x_l^{SSB} < x_l^{M} < x_l^{DFB}$ if $\phi'' < 0$. Proofs have been omitted for brevity and are available upon request.

We can now write the constraints for the second best problem as follows:

$$\begin{aligned}
IR_l &: \quad n_l x_l - p_l + \frac{\Delta\phi(x_l)}{\rho+\delta}(n_h x_h - p_h) \geq 0 \\
IR_h &: \quad n_h x_h - p_h + \frac{\delta}{\rho+\Delta\phi(x_l)}(n_l x_l - p_l) \geq 0 \\
IC_l &: \quad \frac{\rho+\delta}{\rho+\delta+\Delta\phi(x_l)}(n_l x_l - p_l + \frac{\Delta\phi(x_l)}{\rho+\delta}(n_h x_h - p_h)) \geq \\
&\quad\ \frac{\rho+\delta}{\rho+\delta+\Delta\phi(x_h)}(n_l x_h - p_h + \frac{\Delta\phi(x_h)}{\rho+\delta}(n_h x_h - p_h)) \\
IC_h &: \quad \frac{\rho+\Delta\phi(x_l)}{\rho+\Delta\phi(x_l)+\delta}(n_h x_h - p_h + \frac{\delta}{\rho+\Delta\phi(x_l)}(n_l x_l - p_l)) \geq \\
&\quad\ \frac{\rho+\Delta\phi(x_l)}{\rho+\Delta\phi(x_l)+\delta}(n_h x_l - p_l + \frac{\delta}{\rho+\Delta\phi(x_l)}(n_l x_l - p_l))
\end{aligned}$$

IC_h can be simplified to:

$$IC_h' : n_h x_h - p_h \geq \ n_h x_l - p_l$$

which is the same as the incentive compatibility constraint for high types when consumers are myopic. We can further simplify the problem by rewriting IR_h as:

$$n_l x_l - p_l + \frac{\rho+\Delta\phi(x_l)}{\delta}(n_h x_h - p_h) \geq 0$$

Now, comparing IR_l and IR_h, if in the optimum $n_h x_h - p_h > 0$, then IR_l implies that IR_h is satisfied. In IR_h, either $n_l x_l - p_l$ or $n_h x_h - p_h$ must be positive (or both). From IC_h' we know that at least $n_h x_h - p_h$ must be positive.[14] Therefore, we need no longer consider IR_h. Second, it can't be that both constraints don't bind, because the monopolist could raise both p_h and p_l by ε, satisfy the incentive constraints and increase profits. Therefore IR_l must bind. Using this, we find properties of the solution:

Proposition 7 *The solution x_l^*, x_h^* satisfies the following properties:*
i) IC_h' binds and IC_l does not bind,
ii) no distortion at the top: $n_h - c'(x_h^*) = 0$,
iii) more forward looking behavior means higher x_l^* and higher p_l^*: $\frac{dx_l^*}{d\Delta} > 0$ and $\frac{dp_l^*}{d\Delta} > 0$ and
iv) the effect on p_h^* is ambiguous:

$$\frac{dp_h^*}{d\Delta} = \frac{(\rho+\delta)(n_h - n_l)}{(\rho+\delta+\Delta\phi(x_l))^2}(\phi(x_l)x_l - \frac{dx_l}{d\Delta}(\rho+\delta+\Delta(\phi-\phi' x_l))).$$

[14] Suppose not, then IC_h' implies

$$0 > n_h x_h - p_h \geq n_h x_l - p_l > n_l x_l - p_l$$

contradicting that either $n_l x_l - p_l$ or $n_h x_h - p_h$ must be positive.

The proof is in the appendix. The structure of the solution is similar to the continuous model: IR_l and IC_h are binding and there is no distortion at the top.

We are now able to compare the myopic and forward looking solutions. As consumers better understand (higher Δ) the upgrading mechanism (to a type n_h that enjoys a strictly positive surplus: $n_h x_h - p_h > 0$), they are willing to consume more to try to become a high type. Moreover, they are also willing to pay more for the product. This is in line with the intuition that the forward looking low valuation types now value consumption more and have their incentives closer aligned with the monopolist.

Finally, the reason why the effect of Δ on p_h is ambiguous becomes clear when considering the high type's incentive compatibility constraint. To the extent that Δ raises p_l, p_h can increase as well without inducing the high type to switch to the low type's option. However, Δ raises x_l as well, which makes mimicking the low type more attractive. Hence p_h needs to fall to prevent that.

4.2.2. Harmful Addiction

In this section, we examine the impact on the monopolist's decision when consumers are negatively impacted by being addicted and consuming large amounts.[15] We model this in the simplest way possible, setting the cost of addiction as a fixed disutility D per period for the high type.[16] We also define B as the present discounted value of lifetime utility when not using the product. This makes the outside option of the consumer a parameter.

The expressions for V_l, V_h now become:

$$
\begin{aligned}
V_l &= \frac{1}{\rho + \Delta\phi(x_l) - \frac{\delta\Delta\phi(x_l)}{\rho+\delta}}\left(n_l x_l - p_l + \frac{\Delta\phi(x_l)}{\rho+\delta}(n_h x_h - p_h - D)\right) \\
V_h &= \frac{1}{\rho + \delta - \frac{\delta\Delta\phi(x_l)}{\rho+\Delta\phi(x_l)}}\left(n_h x_h - p_h - D + \frac{\delta}{\rho+\Delta\phi(x_l)}(n_l x_l - p_l)\right)
\end{aligned}
$$

To define the individual rationality constraint for the high type, we need to know what happens if the high type quits using the product forever. Once consumption stops, the person stays as a high type for a while, until he transitions[17] into not being an addict and not using the product.

$$\rho V^h_{quit} = -D + \delta(B - V^h_{quit})$$

Hence, the individual rationality constraint for the high type can be written as

$$V_h \geq \frac{-D + \delta B}{\rho + \delta}$$

[15]As we note in the literature review, this section brings us closer to the rational addiction model. The negative impact is not modelled as the effect of a stock variable, but the benefit from this is that we are able to discuss price discrimination among consumers who are heterogeneous in their valuations. We are also able to talk about how consumers' own understanding of their ability to become a higher type (more addicted) affects the solution.

[16]If we had allowed D to vary with the consumption x, this would effectively reduce n_h, a comparative static that we can analyze using the results of the previous section.

[17]We assume the transition probability from going cold turkey to not using the product is the same as transitioning from being a high type to low type for notational convenience. In practice, they can surely differ.

Together with $V_l \geq B$, we get the following two IR constraints:

$$
\begin{aligned}
(IR_h) \qquad & \frac{\rho+\Delta\phi}{\delta}(n_h x_h - p_h) + (n_l x_l - p_l) & \geq & \ (\rho + \frac{\rho\Delta\phi}{\rho+\delta})B + D\frac{\Delta\phi}{\rho+\delta} \\
(IR_l) \qquad & \frac{\Delta\phi}{\rho+\delta}(n_h x_h - p_h) + (n_l x_l - p_l) & \geq & \ (\rho + \frac{\rho\Delta\phi}{\rho+\delta})B + D\frac{\Delta\phi}{\rho+\delta}
\end{aligned}
$$

Note that IC for the high type is not affected:

$$n_h x_h - p_h - D \geq n_h x_l - p_l - D \tag{18}$$

Using the same methods as in the previous section, we can prove (i) that IR_l and IC_h imply that IR_h doesn't bind, (ii) IR_l and IC_h do bind, and (iii) at the optimum, IC_l doesn't bind.[18]

With IC_h and IR_l binding, prices are:

$$p_h = n_h x_h - \frac{1}{\rho+\delta+\phi_c\phi(x_l)}((\rho+\delta)(n_h-n_l)x_l + \Delta\phi(x_l)D) - \rho B \tag{19}$$

$$p_l = n_l x_l + \frac{\phi_c\phi(x_l)}{\rho+\delta+\phi_c\phi(x_l)}((n_h-n_l)x_l - D) - \rho B \tag{20}$$

The optimization problem can be written as:

$$
\begin{aligned}
\max_{x_h,x_l} \qquad & \frac{1}{\delta+\phi(x_l)}(\delta(n_l x_l - c(x_l)) + \phi(x_l)(n_h x_h - c(x_h))) - (n_h - n_l)x_l\frac{\phi(x_l)}{\delta+\phi(x_l)}\frac{\rho+(1-\Delta)\delta}{\rho+\delta+\Delta\phi(x_l)} \\
- \quad & D\frac{\Delta\phi(x_l)}{\rho+\delta+\Delta\phi(x_l)} - \rho B
\end{aligned} \tag{21}
$$

Once again, x_h is the efficient one. The first order condition for x_l can be written as

$$
\begin{aligned}
& \frac{\delta}{\delta+\phi}(n_l - c'(x_l)) + \frac{\delta\phi'}{(\delta+\phi)^2}(n_h x_h - c(x_h) - (n_l x_l - c(x_l))) = \\
& \frac{n_h-n_l}{\delta+\phi}\frac{\rho+\delta(1-\Delta)}{\rho+\delta+\Delta\phi}(\phi + \frac{\delta\phi' x_l}{\delta+\phi} - \phi x_l\frac{\Delta\phi'}{\rho+\delta+\Delta\phi}) + D\frac{\rho+\delta}{(\rho+\delta+\Delta\phi)^2}\phi'(x_l)\Delta
\end{aligned} \tag{22}
$$

We summarize the properties of the solution as follows:

Proposition 8 *The solution x_l^{**}, x_h^{**} satisfies the following properties:*
i) IC_h binds and IC_l does not bind,
ii) no distortion at the top: $n_h - c'(x_h^{**}) = 0$,
iii) when addiction is costlier, consumption for the low type decreases: $\frac{dx_l^{**}}{dD} < 0$
iv) when addiction is costlier and $(n_h - n_l)x_l^{**} - D > 0$, p_l^{**} decreases ($\frac{dp_l^{**}}{dD} < 0$), otherwise the effect is ambiguous
v) the effect of D on p_h^{**} is ambiguous
vi) an increase in the benefits to quitting, B, doesn't change x_l^{**}, but decreases both p_l^{**} and p_h^{**}

[18] These proofs are available upon request.

We do not prove these results here, as they are found in a similar way to Proposition 7.

Costly addiction lowers the quality that low types consume. High types still consume the same quality, as they are not affected on the margin. The price changes depend on the parameters, and can't be signed.

We can also examine how profits for the monopolist change as consumers become more forward looking in this environment. We know:

$$\frac{d\pi}{d\Delta} = \frac{\partial\pi}{\partial x_l}\frac{\partial x_l}{\partial\Delta} + \frac{\partial\pi}{\partial x_h}\frac{\partial x_h}{\partial\Delta} + \frac{\partial\pi}{\partial\Delta} = \frac{\partial\pi}{\partial\Delta}$$

from the envelope theorem. Therefore,

$$\frac{d\pi}{d\Delta} = ((n_h - n_l)x_l - D)\frac{\phi(\rho+\delta)}{(\rho+\delta+\Delta\phi)^2}$$

Clearly, when $D = 0$, we get the same result as before, that when consumers are more forward looking, profits increase. However, D adds a countervailing effect and could reverse this.

5. Conclusion

This paper develops a framework for analyzing dynamic adverse selection problems where types may change over time depending on the schedule the principal offers. The key element used in order to attain tractability is that agents are anonymous, which is often the case. We apply the model to monopoly pricing, analyzing price discrimination for goods where current consumption affects future valuations. We find that the monopolist will increase quality for low types to encourage type upgrading. It may offer substantial discounts at a loss in order to reinforce this effect. For forward looking consumers, the adverse selection effect distorts quality upwards in contrast to the traditional static (and myopic) case. Moreover, the monopolist has incentives to make consumers forward looking as long as the good is not harmful, so that it need not sacrifice rents to encourage upgrading.

The methodology we develop here could be useful for other dynamic mechanism design problems where learning by doing plays an important role. Some natural extensions of the model include the regulation of a firm where marginal costs decrease with know-how and previous production or repeated auctions where consumers learn their demand for the good (as in electricity sales).

6. Appendix

6.1. Solving for F(n) and $\bar{n}$ in pricing below cost example

To determine $F(n)$ we assume that $\alpha(n) = \alpha \cdot (n_1 - n)$ for some scalar $\alpha > 0$. Substituting this into our steady state equation, we find

$$F'(n) = \frac{\delta}{\phi\alpha(n_1 - n)}(G(n) - F(n))$$

For $n > n_{1g}$, $G(n) = 1$. Using $F(n_1) = 1$ we find that

$$F(n) = 1 - C(\frac{\alpha\phi}{\delta}(n_1 - n))^{\frac{\delta}{\alpha\phi}} \text{ for } n > n_{1g}$$

for some constant $C > 0$ to be determined below.

For $n < n_{1g}$, $G(n) = \frac{n-n_0}{n_{1g}-n_0}$, yielding

$$F(n) = \frac{\frac{\alpha\phi}{\delta}(n_1 - n_0) - (n - n_0) + \frac{\alpha\phi}{\delta}(n_1 - n)^{\frac{\delta}{\alpha\phi}}(n_1 - n_0)^{1-\frac{\delta}{\alpha\phi}}}{(\frac{\alpha\phi}{\delta} - 1)(n_{1g} - n_0)} \text{ for } n \leq n_{1g}$$

The constant C is found by the continuity of $F(.)$ at $n = n_{1g}$. It turns out that

$$C = \frac{(\frac{\alpha\phi}{\delta})^{1-\frac{\delta}{\alpha\phi}}}{\frac{\alpha\phi}{\delta} - 1} \frac{(n_1 - n_0)^{1-\frac{\delta}{\alpha\phi}} - (n_1 - n_{1g})^{1-\frac{\delta}{\alpha\phi}}}{n_{1g} - n_0}$$

Since $\bar{x} = 2n_{1g}$ we can now solve for $\bar{n} > n_{1g}$. We know that for $n > n_{1g}$ it is the case that $\frac{1-F(n)}{f(n)} = \frac{\alpha\phi}{\delta}(n_1 - n)$. Thus we find

$$\bar{n} = \frac{\frac{\alpha\phi}{\delta}n_1 + 2n_{1g}}{1 + \frac{\alpha\phi}{\delta}}$$

6.2. Proof of Proposition 4 ($V'(n) \geq 0$)

Consider two types $n' > n$ where n' can mimic n. We see that

$$\begin{aligned}(\rho + \delta)V(n') &\geq n'x(n) - p(n) + \alpha(n')\phi(x(n))V'(n') + \delta \int_{n_0}^{n_1} V(m)g(m)dm \\ (\rho + \delta)V(n) &= nx(n) - p(n) + \alpha(n)\phi(x(n))V'(n) + \delta \int_{n_0}^{n_1} V(m)g(m)dm\end{aligned}$$

Subtract these two equations and consider the case where $n' = n_1$ (and thus $\alpha(n_1) = 0$):

$$(\rho + \delta)(V(n_1) - V(n)) \geq (n_1 - n)x(n) - \phi(x(n))V'(n)\alpha(n)$$

If there exists n such that $V'(n) < 0$ then this equation implies that $V(n_1) > V(n)$. Hence there must exist $n'' \in \langle n, n_1 \rangle$ such that $V'(n'') > 0$. Since $V'(n) < 0$ and $V'(n'') > 0$ it must be the case that over some range $V''(.) > 0$. In particular, there must exist $\tilde{n}$ such that $V'(\tilde{n}) < 0$ and $V''(\tilde{n}) > 0$. However, this contradicts equation (14). Hence $V'(n) \geq 0$ for all $n \in [n_0, n_1]$.
Q.E.D.

6.3. Characterization of the Forward Looking solution

We first prove a corollary to proposition 4:

Corollary 1 *For functions $V(.)$ that are twice differentiable we find that there exists $c > 0$ such that $V'(n) \geq c$.*

Proof: Consider the differential equation (which follows from (14))

$$V''(n) - \beta(n)V'(n) = -\gamma(n)$$

where

$$\begin{aligned}\beta(n) &= \frac{\rho+\delta-\alpha'(n)\phi(x(n))}{\alpha(n)\phi(x(n))} > 0\\ \gamma(n) &= \frac{x(n)}{\alpha(n)\phi(x(n))} > 0\end{aligned}$$

The solution can then be written as

$$V'(n) = \left(C - \int_{n_0}^{n} \gamma(t)e^{-\int_{n_0}^{t}\beta(s)ds}dt\right)e^{\int_{n_0}^{n}\beta(t)dt}$$

Next note that $\alpha(n_1) = 0$ implies that $V'(n_1) = \frac{x(n_1)}{\rho+\delta-\alpha'(n_1)\phi(x(n_1))} > 0$. Because $\gamma(n) > 0$, if $V'(n)$ would become non-positive, it would happen at n_1, but it does not. As $V'(n)$ is strictly positive for n_1, it is strictly positive for all n. The other terms are all nonzero and finite and hence we find $V'(n) \geq c$ for some $c > 0$. (Note, however, that it does not follow that $V'(n)$ is decreasing in n). *Q.E.D.*

The other necessary conditions for the solution to the forward looking case (besides equation (15) in the text) are:

$$\begin{aligned}\lambda'(n) &= \rho\nu(n) + \mu(n)\delta && (23)\\ 0 &= nx(n) - c(x(n)) - \lambda(n) + \mu(n)\alpha(n)\phi(x(n)) && (24)\\ \psi'(n) &= -\rho(1-F(n)) + \xi(n)(\rho+\delta-\alpha'(n)\phi(x(n))) && (25)\\ 0 &= \psi(n) + \xi(n)\alpha(n)\phi(x(n)) && (26)\end{aligned}$$

Solving the fifth equation for $\xi(n) = -\psi(n)/(\alpha(n)\phi(x(n)))$ and plugging that into the fourth equation gives us a differential equation for $\psi(n)$. Now we prove that the solution must have the constraint $V'(n) \geq c$ binding only on an interval that begins at n_0 (or may only bind at n_0).[19] This allows us to conclude that $\psi(n) \geq 0$ for all n.

Proposition 9 *The term $\psi(n) \geq 0$ for all n.*

Proof I. There is no solution where $\psi(n_0) = \psi(n_1) = 0$ and $V'(n) > c$ for all n
Proof by contradiction: assume $\psi(n_0) = \psi(n_1) = 0$ and $V'(n) > c$ for all n. Then:

$$\psi(n) = (K + \int_{n}^{n_1} \rho(1-F(s))e^{\int_{n_0}^{s}\frac{\rho+\delta-\alpha'\phi(x)}{\alpha\phi(x)}du}ds)e^{-\int_{n_0}^{n}\frac{\rho+\delta-\alpha'\phi(x)}{\alpha\phi(x)}ds}$$

[19] This proof uses the necessary conditions that (i) $\psi(n_0) = 0$ if $V'(n_0) > c$, and $\psi(n_0) \geq 0$ if $V'(n_0) = c$ and (ii) $\psi(n_1) = 0$ if $V'(n_1) > c$, and $\psi(n_1) \leq 0$ if $V'(n_0) = c$. A derivation of these conditions can be found in Kamien and Schwartz (1991).

where K is the constant from integration. From $\psi(n_1) = 0$, we have $K = 0$ and $\psi(n) \geq 0$ for all n.

From $\psi(n_0) = 0$, we have $\int_{n_0}^{n_1} \rho(1 - F(s))e^{\int_{n_0}^{s} \frac{\rho+\delta-\alpha'\phi(x)}{\alpha\phi(x)}du} ds = 0$. Since the integral must be positive, this is a contradiction.

II. It can't be the case that $\psi(n_1) < 0$ and $V'(n_1) = c$.

Proof by contradiction: suppose $\psi(n_1) < 0$. Using equation (26) and the fact that $\alpha(n_1) = 0$, $\xi(n_1) \to \infty$. But if $\xi(n_1) \to \infty$, then equation (15) indicates that $x(n_1) \to -\infty$, which cannot be optimal.

III. It can't be the case that $V'(n) = c$ for an interior interval $[n_a, n_b]$ where $n_a > n_0$.

If this was the case, $V'(n_a) = c$ and it must be that $V''(n_a - \varepsilon) < 0$ where ε is small. From the incentive compatibility equation (14), however, $V''(n_a - \varepsilon) < 0$ implies $V'(n_a - \varepsilon) < c$, which is a contradiction.

IV. The only case remaining is where $\psi(n_0) > 0$ and $V'(n_0) = c$. *Q.E.D.*

The proposition then implies that $\xi(n) \leq 0$ for all n. In the next corollary we prove that $\xi(n_1) = 0$.

Corollary 2 *The term $\xi(n) \leq 0$ for all n and $\xi(n_1) = 0$.*

Proof Rewriting equation (26) and taking the derivative with respect to n gives us $\psi'(n) = -\xi'(n)\alpha(n)\phi(x(n)) - \xi(n)\alpha'(n)\phi(x(n)) - \xi(n)\alpha(n)\phi'(x(n))x'(n)$. Evaluating at $n = n_1$ and comparing to equation (25) evaluated at the same point proves that $\xi(n_1) = 0$. *Q.E.D.*

6.4. Proof of Proposition 6

We denoted the solution when the consumer is myopic (and the firm forward looking) by $(x^M(.), p^M(.))$. Let $(x^F(.), p^F(.))$ denote the solution with forward looking consumers (and forward looking firms). We are going to prove that profits from the F case are larger than profits from the M case. We do this in two parts:

I. Downward Deviations (a type n doesn't want to pretend to be a lower type)

We start by fixing the solution to the M case $(x^M(.), p^M(.))$ and showing that a schedule with slightly higher prices $(x^M(.), p^M(.) + \varepsilon)$ satisfies the constraints of the F case.

From the value function for a forward looking customer:

$$(\rho + \delta)V(n_0) = \underbrace{n_0 x^M(n_0) - p^M(n_0)}_{=0} + \alpha(n_0)\phi(x^M(n_0))V'(n_0) + \delta \int_{n_0}^{n_1} V(m)g(m)dm$$

It therefore follows that $V(n_0) > 0$ and there exists an $\varepsilon > 0$ such that the pricing schedule $p^M + \varepsilon$ still satisfies the IR constraint.

Next, consider the IC constraint. We first argue that a type n who is offered the menu $(x^M(n), p^M(n) + \varepsilon)$ will not choose $(x^M(n'), p^M(n') + \varepsilon)$ for some $n' < n$. It is straightforward to see that

$$nx^M(n) - p^M(n) \geq nx^M(n') - p^M(n') \quad \forall n' < n$$

implies that (using the fact that $dx^M(n)/dn \geq 0$):

$$nx^M(n) - p^M(n) - \varepsilon + \alpha(n)\phi(x(n))V'(n) + \delta\int_{n_0}^{n_1} V(m)g(m)dm \geq$$

$$nx^M(n') - p^M(n') - \varepsilon + \alpha(n)\phi(x(n'))V'(n) + \delta\int_{n_0}^{n_1} V(m)g(m)dm \quad \forall n' < n$$

II. Upward Deviations (if an n type pretends to be a higher type, the monopolist makes larger profits)

Here we prove that the firm can only profit from a customer overstating his type, *i.e.* if a type n chooses $(x^M(n'), p^M(n') + \varepsilon)$ rather than $(x^M(n), p^M(n) + \varepsilon)$, where $n < n'$ the firm makes larger profits. Hence the profits in the F case are at least $\varepsilon > 0$ higher than in the M case.

Proposition 10 $\pi'(n) \geq 0$.

Proof Suppose it is not true, *i.e.* there exists $\tilde{n}$ such that $\pi'(\tilde{n}) < 0$. Then one of two cases must occur:

either (I) $\pi'(n) \leq 0$ for all $n \in [\tilde{n}, n_1]$ or (II) there exists $n' \in \langle \tilde{n}, n_1]$ such that $\pi'(n') > 0$.

We argue that neither of these two cases can happen.

First, consider case (I). In this case, the profits of the firm are higher if all types $n \geq \tilde{n}$ are offered $p(\tilde{n}), x(\tilde{n})$ since in that case profits are constant per type (for $n \geq \tilde{n}$) and hence $\pi'(n) = 0$ for all $n \in [\tilde{n}, n_1]$ which implies higher profits than the original menu $(p(n), x(n))$ which featured $\pi'(n) < 0$.

Second, consider case (II). Since $\pi'(\tilde{n}) < 0 < \pi'(n')$ there exist values for n such that $\pi'(n) \leq 0$ and $\pi''(n) > 0$. Consider two values $n^* < n^{**}$ which are close together and satisfy $\pi'(n^*) < 0 = \pi'(n^{**})$. As the two points are close together, one option for the firm is to offer $p(n^*), x(n^*)$ to type n^{**} as well, but presumably the firm can do better. Thus we have

$$(\rho + \delta)\pi(n^{**}) \geq p(n^*) - c(x(n^*)) + \delta\int_{n_0}^{n_1} \pi(m)g(m)dm >$$

$$p(n^*) - c(x(n^*)) + \phi(x(n^*))\alpha(n^*)\pi'(n^*) + \delta\int_{n_0}^{n_1} \pi(m)g(m)dm = (\rho + \delta)\pi(n^*)$$

However, this contradicts $\pi'(n^*) < 0$ as $\pi(n^{**}) > \pi(n^*)$.

Thus it must be the case that $\pi'(n) \geq 0$. *Q.E.D.*

6.5. Proof of Proposition 7

We first assume that (IR_l) and (IC_h) bind while (IC_l) is not binding. At the end of the proof below we verify that in our solution (IC_l) is indeed not binding. Solving the firm's optimization problem with p_l solved from IR_l we get:

$$\max_{p_h, x_l, x_h} \frac{1}{\rho}\left(\frac{\delta}{\delta + \Delta\phi(x_l)}(n_l x_l + \frac{\Delta\phi(x_l)}{\rho + \delta}(n_h x_h - p_h) - c(x_l)) + \frac{\Delta\phi(x_l)}{\delta + \Delta\phi(x_l)}(p_h - c(x_h))\right)$$

$$s.t. \quad IC_h' \quad : \quad \left(\frac{\rho + \delta + \Delta\phi(x_l)}{\rho + \delta}\right)(n_h x_h - p_h) \geq (n_h - n_l)x_l$$

Assigning a Lagrange multiplier of μ_h to IC'_h, the first order condition with respect to p_h is:

$$\frac{\rho\Delta\phi(x_l)}{(\rho+\delta)(\delta+\Delta\phi(x_l))} - \mu_h(\frac{\rho+\delta+\Delta\phi(x_l)}{\rho+\delta}) = 0$$

Hence $\mu_h = \frac{\rho\Delta\phi(x_l)}{(\rho+\delta+\Delta\phi(x_l))(\delta+\Delta\phi(x_l))}$. The first order condition with respect to x_h can be written as

$$\frac{\Delta\phi(x_l)}{(\delta+\Delta\phi(x_l))}(n_h - c'(x_h)) = 0$$

This gives efficiency for the high type, or $n_h = c'(x_h)$.

The first order condition for x_l can be written as

$$\begin{aligned}&\frac{\delta}{\delta+\phi}(n_l - c'(x_l)) + \frac{\delta\phi'}{(\delta+\phi)^2}(n_h x_h - c(x_h) - (n_l x_l - c(x_l))) - \\ &\frac{n_h - n_l}{\delta+\phi}\frac{\rho+\delta(1-\Delta)}{\rho+\delta+\Delta\phi}(\phi + \frac{\delta\phi' x_l}{\delta+\phi} - \phi x_l \frac{\Delta\phi'}{\rho+\delta+\Delta\phi}) = 0\end{aligned} \tag{27}$$

The first line of this equation can be seen as the marginal benefit of raising x_l and the second line as the marginal cost. For existence of an interior optimum we need to assume the first order condition is decreasing in x_l. Therefore, when trying to find $\frac{dx_l^*}{d\Delta}$, we can use the implicit function theorem and sign the derivative with respect to x_l. We will now demonstrate that Δ reduces the second line and hence the marginal cost. This implies that higher Δ leads to higher x_l.

The sign of $d(\text{marginal cost})/d\Delta$ is given by the sign of

$$\begin{aligned}&-\frac{\delta(\rho+\delta+\Delta\phi) + \Delta(\rho+(1-\Delta)\delta)}{(\rho+\delta+\Delta\phi)^2}(\phi + \frac{\delta\phi' x_l}{\delta+\phi} - \phi x_l \frac{\Delta\phi'}{\rho+\delta+\Delta\phi}) \\ &-\frac{\rho+(1-\Delta)\delta}{\rho+\delta+\Delta\phi}\phi x_l \phi' \frac{\rho+\delta}{(\rho+\delta+\Delta\phi)^2}\end{aligned}$$

A sufficient condition for this to be negative is that

$$\phi + \frac{\delta\phi' x_l}{\delta+\phi} - \phi x_l \frac{\Delta\phi'}{\rho+\delta+\Delta\phi} > 0 \tag{28}$$

which can be rewritten as

$$\frac{\delta\phi' x_l}{\delta+\phi} + \frac{\phi}{\rho+\delta+\Delta\phi}(\rho+\delta+\Delta\phi - \phi'\Delta x_l) > 0 \tag{29}$$

Concavity of ϕ and $\phi(0), \phi' \geq 0$ implies that $\phi \geq \phi' x_l$. Hence this inequality is satisfied.

The effect of Δ on p_l is given by

$$\frac{dp_l}{d\Delta} = \left[n_l + \frac{\Delta\phi}{\rho+\delta+\Delta\phi}(n_h - n_l) + (n_h - n_l)x_l \frac{\Delta\phi'(\rho+\delta)}{(\rho+\delta+\Delta\phi)^2}\right]\frac{dx_l}{d\Delta} + (n_h - n_l)x_l \frac{(\rho+\delta)\phi}{(\rho+\delta+\Delta\phi)^2} > 0$$

A similar expression can be derived for $dp_h/d\Delta$ and is given in the proposition.

Finally, we need to verify that in our solution (IC_l) is indeed not binding, as we claim. Note that because of IR_l (binding), IC_l is satisfied if

$$n_l x_h - p_h + \frac{\Delta\phi(x_h)}{\rho+\delta}(n_h x_h - p_h) < 0$$

Using IR_l, this can be written as

$$n_l x_h - p_h < \frac{\phi(x_h)}{\phi(x_l)}(n_l x_l - p_l)$$

Solving IR_l and IC_h for prices yields

$$p_h = n_h x_h - \frac{\rho+\delta}{\rho+\delta+\Delta\phi(x_l)}(n_h - n_l)x_l \quad (30)$$

$$p_l = n_l x_l + \frac{\Delta\phi(x_l)}{\rho+\delta+\Delta\phi(x_l)}(n_h - n_l)x_l \quad (31)$$

Using this, we can rewrite the inequality as

$$\frac{x_h}{\rho+\delta+\Delta\phi(x_h)} > \frac{x_l}{\rho+\delta+\Delta\phi(x_l)}$$

which is true if and only if $x_h > x_l$ (note that the fraction is strictly increasing in x because $\phi(x) > \phi'(x)x$). Finally, we show that equation (27) implies that it is optimal to reduce x_l evaluated at $x_l = x_h$. A sufficient condition for this is

$$n_l - c'(x_h) + \frac{\phi'(x_h)}{\delta+\phi(x_h)}(n_h x_h - n_l x_h) < 0$$

or equivalently $\phi'(x_h)x_h < \delta + \phi(x_h)$ which is indeed satisfied. *Q.E.D.*

References

[1] Atkeson, A. and Lucas, R., "Efficiency and Inequality in a Simple Model of Efficient Unemployment Insurance," *Journal of Economic Theory* 66 (1995), 64-88.

[2] Baron, D, and Besanko, D., "Regulation and Information in a Continuing Relationship," *Information Economics and Policy* 1 (1984), 267-302.

[3] Becker, G. S. and Murphy, K. M., "A Theory of Rational Addiction," *Journal of Political Economy* 96 (1988), 675-700.

[4] Becker, G.S., Grossman, M., and Murphy, K.M., "An Empirical Analysis of Cigarette Addiction," NBER Working Paper #3322 (1994).

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

[6] Driskill, R. and McCafferty, S., “Monopoly and Oligopoly Provision of Addictive Goods,” *International Economic Review* 42 (2001), pp. 43-72

[7] Fethke, G. and Jagannathan, R., “Habit Persistence, Heterogeneous Tastes, and Imperfect Competition,” *Journal of Economic Dynamics Control* 20 (1996), 1193-1207.

[8] Freixas, X., Guesnerie, R. and Tirole, J., “Planning under Incomplete Information and the Ratchet Effect,” *Review of Economic Studies* 52 (1985), 173-192.

[9] Fudenberg, D. and Tirole, J. *Game Theory* (MIT Press, 1991)

[10] Gabszewicz, J., Pepall, L., and Thisse, J-F. “Sequential Entry with Brand Loyalty caused by Consumer Learning-by-Using,” *Journal of Industrial Economics* 40 (1992), 397-416.

[11] Gärtner, D. “Monopolistic Screening Under Learning By Doing”, University of Zurich, SOI Working Paper No. 0718, 2007.

[12] Green E. “Lending and the Smoothing of Uninsurable Income,” in Prescott, E. and Wallace, N. (editors) *Contractual Arrangements for International Trade*. University of Minnesota Press, 1987 .

[13] Kamien, M.I., and Schwartz, N.L. *Dynamic Optimization.* (Amsterdam: North Holland, 1991).

[14] Keller, K.L. *Strategic Brand Management* (New Jersey: Pearson Education, 2003).

[15] Kocherlakota, Narayana, “Advances in Dynamic Optimal Taxation,” University of Minnesota mimeo 2005.

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

[17] Lewis, T. R., and Yildirim, H., “Learning by doing and dynamic regulation,” *RAND Journal of Economics* 33 (2002), 22-36

[18] Schlee, E. “Buyer experimentation and introductory pricing,” *Journal of Economic Behavior & Organization* 44 (2001), 347-362

[19] Shapiro, C. “Optimal Pricing of Experience Goods,” *Bell Journal of Economics* 14 (1983), 497-507.

[20] Shapiro, C.and Varian, H. *Information Rules: A Strategic Guide to the Network Economy* (Cambridge: Harvard Business School Press, 1998).

[21] Showalter, M., “Firm Behavior in a Market with Addiction: The Case of Cigarettes,” *Journal of Health Economics* 18 (1999), 409-427.

[22] Skreta, V., “Sequentially Optimal Mechanisms,” *Review of Economic Studies* 73:4 (2006), 1085-1111.

[23] Thomas, J. and Worrall, T., "Income Fluctuations and Asymmetric Information: An Example of a Repeated Principal-Agent Problem," *Journal of Economic Theory* 51 (1990), 367-390.

[24] Townsend, R.M. "Optimal Multiperiod Contracts and the Gain from Enduring Relationships under Private Information," *Journal of Political Economy* 90 (1982), 1166-1186.