# The Optimal Pricing Scheme when Consumer is Habit Forming

Eleftheria Triviza*

March 7, 2017

**Abstract**

This article analyses how consumers' habit formation affects firms pricing policy. Sophisticated and naive consumers are considered. The former realize that their current consumption will affect future consumption, the latter do not. Our main result, both with symmetric and asymmetric information, is that under naive habit formation, the optimal pricing pattern is a three part tariff, namely a fixed fee, an amount of units priced below cost and after their end pricing above marginal cost. Different from Grubb (2009), we claim that only one mistake, in our case underestimation of future demand, is sufficient for three part tariff to be optimal.

*Department of Economics. Lancaster University Management School. e.triviza@lancaster.ac.uk

# 1 Introduction

In several markets, a prevalent pricing pattern is that of three-part tariff, which includes a fixed fee, an allowance of free units, and a positive price for additional units beyond the allowance. Examples of this kind of markets are *communication services*, i.e. wireless phone services, internet access, and *subscription services*, i.e. on-line music download, on-line newspaper, data center hosting e.t.c. Such a pricing scheme, which features increasing prices, is hard to explain in standard models of non-linear pricing, which tend to predict that marginal prices should rather be decreasing.

These markets[1] have attracted the attention of psychologists, who have found that the stock of past consumption typically affects the consumption today. That is, preferences exhibit habit formation. Some researchers have conjectured that consumption may be even addictive (Bianchi and Phillips, 2005; Park, 2005; Hooper and Zhou, 2007) although the evidence for this does not seem to be conclusive.

The existence and the implications of habit forming behavior has been studied in a number of different economic environments. Two types of habit forming consumers have been studied and considered in this article. On the one hand, the *Sophisticated (rational) Habit Forming* consumer who is aware that the today consumption affects future consumption (Becker and Murphy, 1988; Constantinides, 1990; Abel, 1990; Campbell and Cochrane, 1999; Jermann, 1998; Boldrin et al., 2001; Carroll et al., 2000; Fuhrer, 2000). On the other hand, the *Naive (myopic) Habit Forming* consumer who recognizes that her current satisfaction depends on past habits, but she neglects the impact of current decisions on her future preferences (Pollak, 1970; Loewenstein et al., 2003). Interestingly, Muellbauer (1988) provides an excellent overview of the two extremes, and concludes that the empirical evidence

[1] Mostly, the markets of internet, cell phone calling and messaging.

seems to be in favor of the presence of myopic habits.

This article claims that the habit forming behavior of the consumers may explain why three part tariff is the optimal pricing policy for the firm. It is shown that with sophisticated habit forming consumers it is optimal to charge a two part tariff. However, naive habit formation makes it optimal to charge three part tariff. Moreover, it is shown that if the consumption choice is considered sequentially within the contract period, the consumer undervaluates the offered contract at the contracting period and underestimates the demand conditional on it being high, which are the main characteristics of a naive habit forming behavior, it is optimal for the firm to offer a three part tariff.

The consumption choice is considered sequentially means that the consumer does not form her expectation for future consumption statically as a total consumption in the beginning of the contract period but she considers dynamically each consumption opportunity. The consumer knows that she will update her consumption strategy for every consumption opportunity given her past consumption and the opportunity cost of foregone future purchases during the contract period.

Furthermore, the fact that the consumer is unaware of her habit forming behavior at the contract period has two basic effects on the optimal pricing. First, she undervalues the offered contract since she cannot foresee that she will value the good more the more she consumes. Therefore, the firm cannot absorb all the consumer surplus at the contract period and thus finds it optimal to distort the marginal pricing. Second, she underestimates the probability of consuming in the future, since her past consumption makes her more prompt than she expects to consume in the future.

Naive habit formation could be viewed as an alternative channel other than overconfidence of Grubb (2009) that induces this kind of pricing scheme. Grubb (2009), after considering a number of alternative explanations, as for example the flat-rate bias, hyperbolic discounting e.t.c, claims that in order for such a pricing scheme to be optimal it is necessary that the consumer overestimates her demand conditional on it being low and underestimates the demand conditional on it being high, behavior which is actually captured by overconfidence. This article shows that both mistakes are not necessary if the elements mentioned before are taken into consideration and it is enough that she underestimates the demand conditional on it being high.

We develop a model where the consumer has two consumption opportunities. Firstly, it is considered the sophisticated habit forming consumer and a monopolist. In this case, both the consumer and the monopolist have the same beliefs about the behavior of the consumer at the contractual stage, when she is called to accept or reject the contract. The optimal pricing scheme in this case is *two part tariffs.*

Moreover, we consider a naive habit forming consumer, namely a consumer who realizes that she is habit forming only after she has consumed. In this case, the monopolist has different prior beliefs from the consumer. The firm being longer in the market can recognize the type of the consumer, thus her habit forming behavior and her naivety. Given the naive habit forming behavior of the consumer, the optimal contract offered by the firm resembles a three part tariff. For low volumes, the price is smaller than the marginal cost, and then for high volumes it becomes bigger than the marginal cost. Moreover, when the marginal cost is low enough the optimal pricing scheme resembles even more the observed one, since low volumes are free of charge and high volumes are charged above marginal cost. Interestingly, when we enrich the model with Hotelling competition, which is an important element since this kind of market are characterized by tough competition, the optimal pricing scheme continues being three part tariffs.

Important element of the model is that the firm cannot absorb all the consumer surplus at the contracting period, because the consumer undervalues it. The undervaluation of the contract has two implications. On the one hand, the consumer is left with a mis-perception rent, namely the difference between the true expected surplus and the perceived one. On the other hand, the firm has the incentive to distort the optimal marginal prices in order to mitigate this undervaluation. The more habit forming is the consumer, namely the more she underestimates the utility gained from the consumption of the good, the more distorted are the optimal marginal prices. This is because the need for mitigation of the undervaluation is bigger. Therefore, the bigger the distortion the smaller is the optimal marginal price for low volumes and the bigger for high ones.

Interestingly, even if the optimal marginal price for low volumes is smaller than the marginal cost or even zero, the naive consumer underconsumes compared to the sophisticated one. This is because, the consumer being forward looking but naive, she takes into consideration the price change at the second unit, but not the future benefit of consuming. The firm finds it optimal to charge such a pricing scheme because even if the second period is the period when the firm could take advantage of the mistaken expectations of the consumer for the probability of consumption; at the same time, it is the period when the consumer cannot foresee the real value of the good. Thus, in order to mitigate the undervaluation, the firm finds it optimal to charge prices such that would decrease the probability of consumption for the naive consumer relative to the one that would be optimal for the sophisticated. In this way, it is less probable to consume at the first period and thus less probable for the firm to be at a situation where the consumer undervalues. Similarly, for high volumes the consumer underconsumes. Thus, the firm by charging this kind of contract exacerbate the mistake the consumer does because of her naivety in order to mitigate the profit losses it has.

In section 4, we consider the model with asymmetric information, where the firm cannot observe the consumer type. I study the optimal screening of habit forming consumers of different degrees of sophistication. My main assumption is that the firm knows that all consumers in the market are habit forming and also knows the expected level of demand, but cannot observe the type of the consumer, sophisticated or naive. Thus, the consumer's private information is actually her level of sophistication.

We contend that the menu of contracts that is frequently observed in a number of markets and comprises both two and three-part tariffs are explained by the existence of consumers of diverse sophistication in the market. We show that the firm offers both a two-part and a three-part tariff so as to screen out sophisticated and naive consumers. Moreover, it is claimed that three part tariff is still the optimal contract for naive consumers even when there is assymetric information.

This article shows that both consumer types are left with a positive rent. In particular, the firm cannot exploit[2] the naivety of naive consumers in the same way as analyzed for the full information case. Moreover, the presence of naive consumers in the market extents a positive externality to the sophisticated consumer. This means that the sophisticated consumers are left an information rent, so they take advantage of their superior information.

One might expect that sophisticated consumes would not have any incentive to mimic the naive ones by choosing a contract that penalizes large consumption levels, as consumers know that they are more likely to consume more in the future. However, the fact that the firm cannot exploit consumers' naivety, and thus must offer to naive consumers contracts

[2]Exploitative in the sense of Eliaz and Spiegler (2006) where "An exploitative contract extracts more than the agent's willingness to pay, from his first-period perspective"

with a lower fixed fee creates an incentive for sophisticated consumers to mimic the naive ones. Thus, in order to screen out the two types the firm must leave an information rent to sophisticated consumers, making sure that her contract are incentive compatible.

One might also expect that the presence of a sophisticated consumer would mitigate the exploitation of the naive consumers. However, the optimal contract for the naive consumer is still three part tariff as with full information and its distortion as a result of asymmetric information makes the naive worst off. The mis-perception rent that they get decreases with respect to the full information case. We show that the mis-perception rent decreases with the marginal prices, and thus the increase in the marginal price in the contract of the naive consumers, due to the asymmetric information and screening, affects negatively such consumers.

Finally, it is shown that the contract intended for the naive consumers exacerbates even more the mistake of underconsumption that she does due to her naivety. The firm finds it optimal to charge a contract that makes consumption for naive relatively less likely. The intuition is firstly that it makes the contract less attractive for the sophisticated consumers but also minimizes the expected consumer surplus that cannot extract with the fixed fee at the contractual stage, as shown also in the analysis of the full information model.

The article proceeds as follows. Section 2 discusses the related literature. Section 3 is devoted to the benchmark model of full information both for a monopolistic and an oligopolistic market. Section 4 discusses the case of assymetric information. Section 5 makes a comparison between the full and incomplete information case and also develops some comparative statics. Finally, Section 6 summarizes and concludes.

## 2 Literature

This article is related to different streams of the literature. First, it is clearly related to models that try to explain the introduction of three part tariffs. Grubb (2009) shows that over-confidence about the precision of the prediction when making difficult forecasts, free disposal and relatively small marginal cost would explain the use of three part tariff. He claims that three part tariff is the optimal pricing scheme when necessarily the behavior of the consumer is characterized by overestimation of the demand given the demand being low and underestimation of the demand given it being high. In our case, we propose a different behavior that could explain this pricing scheme without necessarily both mistakes being present. Moreover, we study an environment where the firm observes the amount actually consumed by the consumer in each period[3] and not only the amount the consumer has bought.

Grubb (2014) shows that inattentive behavior, having similar features to overconfidence, could explain the introduction of three part tariff. Thus, again it claims that both the mistake of overestimation of the demand at the beginning of the contract period and the mistake of underestimation in the end is needed for such a pricing scheme to be optimal. The common element between our model and Grubb (2014) is that we both consider the consumption dynamically within the contract period but we propose different type of behavior.

Eliaz and Spiegler (2008) consider a model where consumers have biased priors, that we do as well, but only two types of ex post demand high or low. The consumers are optimistic and think that the good state is more probable to happen. They describe a situation where consumers are dynamically inconsistent and they under or overestimate average demand. Thus, Eliaz and Spiegler (2008) studies a completely different behavioral bias, having only

[3]Though, we assume that the firm cannot observe all the consumption opportunities of the consumer.

in common the biased priors.

Moreover, it is related to models of non linear pricing. Articles like Mussa and Rosen (1978) and Maskin and Riley (1984) explain contracts with high marginal prices for early units and marginal cost pricing for late units consumed; though, they cannot predict the inverse which is marginal prices below marginal cost at the early stage and an increase in the marginal prices later on.

In particular, we study the optimal pricing scheme when the good is habit forming thus articles that discuss the optimal pricing of habit goods (Nakamura and Steinsson, 2011; Fethke and Jagannathan, 1996) or even addictive (Becker et al., 1991; Driskill and McCafferty, 2001) are connected to our study. But unlike to this kind of literature we consider habit formation and optimal pricing within a contract period when a contract is signed at a zero period and there is no possibility for the firm to renegotiate the price during the contract period.

Moreover, the discussion of a naive habit forming consumer is closely related to a number of articles that consider the optimal non linear pricing induced by various types of consumers' biases or nonstandard preferences.

On the one hand, there are articles discussing biased beliefs, such as naive quasi-hyperbolic discounting for leisure good (DellaVigna and Malmendier, 2004; Eliaz and Spiegler, 2006), naivety about self-control (Esteban et al., 2007; Heidhues and Kőszegi, 2010) and myopia (Gabaix and Laibson, 2006; Miao, 2010). A common result of these articles is that these behavioral biases lead to underestimation of the demand, and hence to marginal prices above marginal cost. Thus, these models cannot explain why marginal prices are below marginal cost for low volumes.

On the other hand, biases like naive quasi-hyperbolic discounting for investment goods

(DellaVigna and Malmendier, 2004) and flat rate bias (Herweg and Mierendorff, 2013; Lambrecht and Skiera, 2006), that lead to overestimation of demand, can explain prices below marginal costs, but not above.

DellaVigna and Malmendier (2004) were the first to point out that firms might design contracts so as to exacerbate consumer's mistakes. Since their pioneering contribution, number of articles have explored specific way of exploiting consumer naivety. In our model, the firm does offer a contract that exacerbates consumer's mistake but cannot extract all the consumer surplus produced.

This article is also related to the literature on exploitative contracting, where firms design their contracts to profit from the agent's mistakes, and other considerations or constraints are non-existent, not binding, or not central. There are two kind of consumers' mistakes that are more often analyzed in the literature. Firstly, the consumer does not understand all features of a contract (all prices and fees) [Gabaix and Laibson (2006), Armstrong and Vickers (2012)]. For example, in Gabaix and Laibson (2006), consumers underestimate the probability of needing an add-on after buying the good. An other kind of mistake is to mispredict the own behavior with respect to the product (DellaVigna and Malmendier, 2004). The latter kind of mistake is closer to the model we study here, as in our model the consumer mispredicts that her valuation for the good will change if she has consumed before.

The section of asymmetric information is clearly related to the behavioral screening literature, where a principal screens the agents with respect to their cognitive features (such as loss aversion [Hahn et al. (2012), Carbajal and Ely (2012)], present bias, temptation disutility [Esteban et al. (2007)], or overconfidence [Sandroni and Squintani (2010), Spinnewijn (2013)]

Rubinstein (1993) studies for the first time the problem of a principal who wishes to discriminate between consumer types according to their cognitive features. In that article, consumers have bounded ability to categorize realizations of a random variable. Different consumer types have different categorization abilities, and the principal's optimal contract is designed to screen their type. Piccione and Rubinstein (2003) perform a similar exercise, when different consumers differ in their ability to perceive temporal patterns.

Eliaz and Spiegler (2006) claim that the ultimate source of gains for the principal is the non-common prior assumption which they interpreted as "a situation in which the agents have a systematic bias in forecasting their future tastes, whereas the principal has an unbiased forecast". This assumption seems to be really important in the case they study where the consumer is uncertain as to whether her preference will change, but knows exactly what they could change into. The firm takes advantage of its superior information and contracts also the event that the consumer thinks unlikely to happen. Even if the consumers they study are dynamically inconsistent and evaluate their future actions according to their first period utility function, the fact that they know what their taste could change into leaves space for exploitation. In our case, in contrast, the consumer does not know that her utility function will change after consuming in the first period, so the firm cannot exploit its superior information that the consumer is habit forming. This feature becomes important because in both cases the contract is singed before the consumer experiences the change in her utility and cannot be renegotiated afterwards.

# 3 Full Information

This section presents the basic structure of the model and the definition of the optimal pricing policy under full information.

The model follows Grubb (2014) in modeling a consumer who has two consumption opportunities one per period and in each period purchases at most 1 unit of the good. Moreover, the consumer is ex-ante uncertain about her per-period evaluation of the service.

### The Model

Consider a model where there are habit forming consumers of different sophistication and one firm. The consumers are uncertain about their valuation of the good in each period.

The time horizon is $T = 2$. At period 0, the firm offers a menu of contracts:

$$\mathbf{p}^\theta = \{F^\theta, p_1^\theta, p_2^\theta\}$$

The contract $\mathbf{p}^\theta$ consist of $p_1^\theta$ (the price of the first unit consumed), $p_2^\theta$ (the price of the second unit consumed), and $F^\theta$ (a fixed payment). The first unit has the same price irrespectively of the period $t$ when is consumed. Time dependent pricing would require that the firm could observe and record the opportunities to consume, if for example the consumer had a direct communication with the firm in every opportunity to consume. Thus, it is a relevant assumption to assume that the firm cannot observe whether the consumer thinks to consume or not[4] At each consecutive period $t \in \{1, 2\}$, the consumer learns the realization of a taste shock $\upsilon_t$, randomly drawn from a cumulative distribution function $F(\upsilon)$ with support [0,1], the same for all types of consumers and for both periods. This is the valuation that a unit of good has in period $t$. Then, given her valuation, she makes a binary quantity choice $q_t = \{0, 1\}$, considering whether or not to purchase the good[5].

The total payment $p^\theta(\mathbf{p}^\theta, \mathbf{q})$

$$p^\theta(\mathbf{q}) = p_1^\theta q_1 + p_1^\theta(1 - q_1)q_2 + p_2^\theta q_1 q_2 + F^\theta,$$

[4]Contrary to Grubb (2014), we assume that the firm cannot observe the period in which the consumption takes place. In the appendix we saw that the result holds also in the case that the firm could observe it.

[5]We assume that the good is indivisible.

is a function of quantity choices $\mathbf{q} = (q_1, q_2)$ and the pricing scheme $\mathbf{p}^\theta = (F, p_1^\theta, p_2^\theta)$. The timing of the game is described from Figure 1.

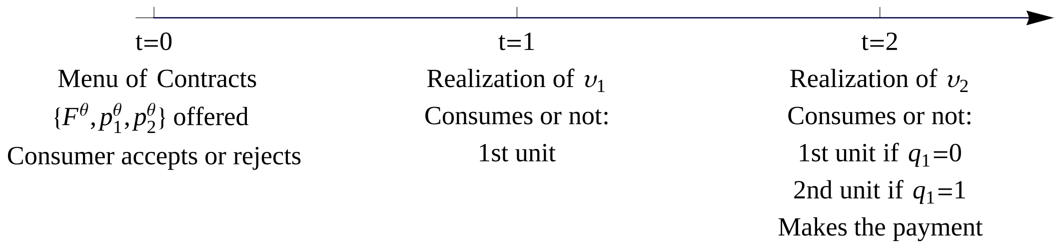



Figure 1: Timing of the game

The optimal consumption strategy, for given prices, is a function mapping valuations to quantities:

$$\mathbf{q}(\boldsymbol{\upsilon}; \mathbf{p}^\theta) : \boldsymbol{\upsilon} \to \mathbf{q}$$

Moreover, the ex ante expected gross utility of the consumer from making optimal consumption choices is:

$$U = E[u(\mathbf{q}(\boldsymbol{\upsilon}; \mathbf{p}^\theta), \boldsymbol{\upsilon}, \mathbf{p}^\theta)]$$

The expected profits per consumer equal the revenues less the variable cost with marginal cost $c \geq 0$ per unit produced. The fixed cost is normalized to zero. Thus, the profit function is:

$$\Pi = E[p^\theta(\mathbf{q}(\boldsymbol{\upsilon}; \mathbf{p}), \mathbf{p}^\theta) - c(q_1(\boldsymbol{\upsilon}; \mathbf{p}^\theta) + q_2(\boldsymbol{\upsilon}; \mathbf{p}^\theta))$$

Finally, the expected social surplus is:

$$S = E[\sum_{t=1}^{2}(v_t - c)\mathbf{q}(\boldsymbol{v};\mathbf{p}^{\theta})]$$

### Consumer

Consider a consumer who is habit forming in the sense that her consumption today is affected by her consumption in previous periods. Her valuation for the service at period t is:

$$\tilde{v}_t = v_t q_t + \theta\beta q_{t-1} q_t$$

This means that if she consumed in the previous period, her valuation for the service today increases by $\theta\beta$, where $0 \leq \beta \leq 1$ is the habit formation coefficient, namely it defines how much habit forming is the consumer, how much she is affected by previous consumption. Moreover, $0 \leq \theta \leq 1$ is the type of the consumer, it is a measure of her naivety and of how much she realizes that she is habit forming. The larger is $\theta$, the less naive is the consumer, so the more fully she realizes that she is affected by her previous consumption. Thus, $\theta = 1$ means that the consumer is sophisticated habit forming, $0 < \theta < 1$ means that she is partially naive, and $\theta = 0$ that she is completely naive.

The expected utility of the consumer at the contracting period is:

$$U(\mathbf{p}) = -F + \int_{v_1^*}^{1}\left(v_1 - p_1 + \int_{p_2-\theta\beta}^{1}(v_2 + \theta\beta - p_2)dF(v_2)\right)dF(v_1) + F(v_1^*)\int_{p_1}^{1}(v_2 - p_1)dF(v_2)$$

The first part is the fixed fee, the second the expected utility if both units are consumed and the third the expected utility if only the second unit is consumed.

Every time that the consumer faces a consumption decision, she compares the valuation of the unit with the reservation price. The reservation price in each period is the optimal threshold above which the valuation should be in order for the consumer to consume the unit. As the valuation of the unit is random and the consumer does not know it ex ante, she calculates the optimal threshold, as an optimal consumption rule different for each potential consumption decision. Thus, these thresholds are the consumption strategy of the consumer and the argument of maximization of her expected utility for the respective unit and period. In this way, the consumer maximizes her ex ante utility.

For simplicity, we assume that there are two types of consumer, a sophisticated habit forming consumer with $\theta = 1$ and a naive one, with $\theta = 0$. In each period, consumers choose the optimal threshold above which is optimal for them to consume. At the contracting period, the consumer does not know the future realizations of her valuation of the good, so chooses which contract to sign on the basis of her expected utility.

**Sophisticated Habit Forming Consumer ($\theta = 1$):** Solving backwards, the second period optimal threshold is, obviously:

$$v_{2S}^* = \begin{cases} p_1 & \text{if} \quad q_1 = 0 \\ p_2 - \beta & \text{if} \quad q_1 = 1 \end{cases}$$

Given $v_{2S}^*$, the first period maximization problem of the sophisticated habit forming consumer is:

$$\begin{aligned} \max_{v_{1S}^*} U^S(\mathbf{p^S}) = & \int_{v_{1S}^*}^1 \left( v_1 - p_1 + \int_{p_2-\beta}^1 (v_2 + \beta - p_2) dF(v_2) \right) dF(v_1) \\ & + F(v_{1S}^*) \int_{p_1}^1 (v_2 - p_1) dF(v_2) - F^S \end{aligned} \tag{1}$$

Maximizing with respect to $v_{1S}^*$ the optimal first period threshold is (after an integration by parts):

$$v_{1S}^* = p_1 - \int_{p_2-\beta}^{p_1} (1 - F(v_2)) dv_2$$

The consumer is forward looking and is aware of being habit forming, so she takes into account both the opportunity cost of consuming the first unit (i.e. the price increase $p_2 - p_1$ for the second unit) and the increase in her valuation due to the habit. The habit forming consumer expects to experience a larger utility in the future, if she consumes the first unit, so she finds it optimal to increase the probability of consuming the first unit. Thus, she decreases the optimal threshold. Moreover, the first period threshold increases if the second unit marginal price increases and decreases if habit formation is stronger.

**Naive Habit Forming Consumer ($\theta = 0$):** In period 2, the optimal threshold is exactly the same as for a sophisticated. However, from the period 1 perspective the consumer anticipates that the second period threshold will be:

$$v_{2N}^* = \begin{cases} p_1 & \text{if} \quad q_1 = 0 \\ \\ p_2 & \text{if} \quad q_1 = 1 \end{cases}$$

That is, the consumer does not anticipate that the first period consumption will affect the valuation of the good in the second period.

Given $v_{2N}^*$, the first period maximization problem of the naive habit forming consumer is:

$$\begin{aligned} \max_{v_{1N}^*} U^N(\mathbf{p^N}) = & \int_{v_{1N}^*}^1 \left( v_1 - p_1 + \int_{p_2}^1 (v_2 - p_2) dF(v_2) \right) dF(v_1) \\ & + F(v_{1N}^*) \int_{p_1}^1 (v_2 - p_1) dF(v_2) - F^N \end{aligned} \tag{2}$$

maximizing with respect to $v_{1N}^*$, the optimal first period threshold is:

$$v_{1N}^* = p_1 + \int_{p_1}^{p_2} (1 - F(v_2)) dv_2$$

Consumer takes into account the opportunity cost of first-period consumption, as she is forward looking, but she does not consider the effect of first period consumption on second-period evaluation. Thus, the optimal threshold is the same as that of a non habit forming consumer. Clearly, $v_{1N}^* > v_{1S}^*$ and $v_{2N}^* > v_{2S}^*$ thus the naive consumer under-consumes in both periods for given marginal prices.

These results follow from the fact that the two types of consumers, sophisticated and naive, are ex post identical; the only difference is the ex ante perception of how much habit forming they are.

In fact, the true ex-ante utility of the consumer is the same as for a sophisticated habit forming consumer, with the difference that the first period optimal threshold is the one of a non habit forming consumer:

$$\begin{aligned}\tilde{U} = &\int_{v_{1N}^*}^1 \left(v_1 - p_1 + \int_{p_2-\beta}^1 (v_2 + \beta - p_2) dF(v_2)\right) dF(v_1) \\ &+ F(v_{1N}^*) \int_{p_1}^1 (v_2 - p_1) dF(v_2) - F^N\end{aligned} \tag{3}$$

Thus, when the consumer chooses the optimal contract, she perceives herself as non habit forming[6]. Therefore, she believes that her expected utility is $U^N$, even though her actual expected utility, and the one that the firm expects that she will have, is $\tilde{U}$.

The consumer tends to consume the first unit less often than she would do if she was aware of her habit forming behavior. In the first period, though, she does not make any mistake as a result of her naivety, in the sense that there is no inconsistency between her previous and her period one self. She uses in period one the same threshold she expected to use when she chose her contract in the contractual stage. The probability of consuming in the first period is $1 - F(v_{1N}^*)$ as it was expected at the period 0. Thus, there is no difference between the consumer expectations about her future self and how she actually acts, namely no mistake that the firm could take advantage of. The only implication of consumer's naivety is related to the expected consumer surplus, which is smaller than the one that would be

[6]This whole analysis holds also when the consumer is partially naive, namely she knows that she is habit forming but she believes that she is less habit forming than she actually is. In this case, the perceived valuation of the good is $\tilde{v}_t = v_t q_t + \hat{\beta} q_{t-1} q_t$ and $\hat{\beta} < \beta$. As in the case of the naive consumer, the partially naive consumer has no mistaken beliefs if the demand is low but she underestimates her demand it is high. See the Appendix.

produced if the consumer was sophisticated. However, this has no direct implication on first unit marginal price as the consumer consumes what she expected[7].

In the second period, given that the consumer has not consumed before ($q_1 = 0$), she does not realize that she is habit forming and thus she consumes as much as she was expecting to consume at the contract period. The probability of consuming is $F(v_{1N}^*)(1 - F(p_1))$ and it is not different from what the consumer would expect. The consumer does not overestimate the probability of buying only one unit, actually does not make any mistake given that her consumption is low.

On the other hand, given that the consumer has consumed before ($q_1 = 1$), she underestimates the probability of consuming two units. She expects that her optimal threshold in this case would be $p_2$ but she realizes later that it is $p_2 - \beta$. Thus, the probability of consuming at the second period is expected to be $(1 - F(v_{1N}^*))(1 - F(p_2))$ but given she has consumed at period 1 it is $(1 - F(v_{1N}^*))(1 - F(p_2 - \beta))$. This follows from the fact that the consumer believes that she is not habit forming and she realizes this only after she has consumed. This means that when she chooses which contract to sign, the consumer underestimates her demand if the demand is high. In other words, she underestimates the probability of consuming the second unit.

**Lemma 1.** *A naive habit forming consumer makes no mistake if her demand is low, but underestimates her future demand when it is high.*

This mistake in forming her expectations is one of the reasons why the consumer undervaluates the offered contract at the contracting stage. Firstly, the consumer does not anticipate that consuming in the first period will increase the valuation of her second unit so she does not expect the $\beta$ additional valuation and she does not consider it in the ex ante valuation of

[7] There is an indirect implication that will become evident at the discussion of firm's maximization problem.

the whole contract. Secondly, as she does not anticipate that she is habit forming she underestimates the probability of consuming $(1-F(v_{1N}^*))(1-F(p_2)) < (1-F(v_{1N}^*))(1-F(p_2-\beta))$ the second unit and thus acquiring this extra utility.

## Firm

Let us consider first the case that the firm can observe the type of the consumer and thus can offer a type specific contract.

### Sophisticated Consumer

There is a monopolistic firm in the market. The cost of production of one unit of the good is $0 \leq c \leq 1$.

The maximization problem of the firm is:

$$\max_{\mathbf{p^S}} \Pi^S = S^S(\mathbf{p^S}) - U^S(\mathbf{p^S}) \quad \text{s.t} \quad U^S(\mathbf{p^S}) \geq 0$$

It is the difference between the expected gross surplus produced minus the expected consumer surplus subject to the participation constraint. The expected gross surplus is:

$$\begin{aligned} S^S(\mathbf{p^S}) = &\int_{v_{1S}^*}^1 \left( v_1 - c + \int_{p_2-\beta}^1 (v_2 + \beta - c) dF(v_2) \right) dF(v_1) \\ &+ F(v_{1S}^*) \int_{p_1}^1 (v_2 - c) dF(v_2) \end{aligned} \tag{4}$$

Maximizing with respect to $\mathbf{p^S}$, the optimal contract is found and it is given by the following Lemma:

**Lemma 2.** *If the consumer is sophisticated habit forming, the equilibrium allocation is the first best allocation. There is marginal cost pricing, namely the prices that maximize the profits of the firm are $(p_1, p_2) = (c, c)$ and the fixed fee, $F^S$, equals to the consumer surplus.*

The firm maximizes its profit by charging marginal prices that induce the first best allocation and then with the fixed fee $F^S$ it absorbs all the consumer surplus. The firm does the rent extraction in such a way that is balanced with the participation as in a basic monopoly pricing problem. (see Appendix)

**Naive Consumer**

Let now consider how the maximization problem of the firm changes when the consumer is naive habit forming.

**Monopoly:** The firm recognizes that it faces a naive habit forming consumer whose participation depends on her mistaken expected utility. Moreover, it knows that the social surplus that is produced is given by:

$$S^N(\mathbf{p^N}) = \int_{v_{1N}^*}^{1} \left( v_1 - c + \int_{p_2-\beta}^{1} (v_2 + \beta - c) dF(v_2) \right) dF(v_1) \\ + F(v_{1N}^*) \int_{p_1}^{1} (v_2 - c) dF(v_2)$$

The firm considers that in the first period the consumer does not know that she is habit forming and consumes only if the valuation of the unit is greater than $v_{1N}^*$. Moreover, it takes into account that given that she has consumed in the first period, her valuation of the good in the second period is higher, since it is affected by past consumption. Therefore, at the contractual stage it takes into consideration that the consumer will update her second unit threshold and her valuation for the second unit, if she has consumed in the first period.

The firm maximizes its profits, which are the difference between the social surplus and the consumer surplus, subject to the participation constraint of the consumer. In this case, though, the true consumer surplus produced $\tilde{U}$ (equation 3) is different from the one the consumer perceives at the contracting period $U^N$ (equation 2). Thus, the optimization problem of the firm is:

$$\begin{aligned}
\max_{U^*,p_1,p_2} \Pi &= S^N(\mathbf{p^N}) - \tilde{U}(\mathbf{p^N}) \\
&= S^N(\mathbf{p^N}) - U^N(\mathbf{p^N}) - (\tilde{U}(\mathbf{p^N}) - U^N(\mathbf{p^N})) \\
&= S^N(\mathbf{p^N}) - U^N(\mathbf{p^N}) - \Delta \\
\text{s.t.} \quad & U^N(\mathbf{p^N}) \geq 0
\end{aligned}$$

where

$$\Delta = \tilde{U}(\mathbf{p^N}) - U^N(\mathbf{p^N})$$

is the difference between the true expected utility from the contract $\tilde{U}(\mathbf{p^N})$ (equation 3) and the perceived utility $U^N(\mathbf{p^N})$ (equation 2). It follows that the firm cannot extract all the consumer surplus. Moreover, the firm chooses a pricing scheme that makes the participation constraint bidding, $U^N(\mathbf{p^N}) = 0$. Still, though there is a positive rent $\Delta$ that must be left to the consumer. After some simplifications, $\Delta$ can be rewritten as:

$$\Delta = (1 - F(v_{1N}^*)) \left( \int_{p_2-\beta}^{p_2} (1 - F(v_2)) dv_2 \right)$$

Then, the maximization problem of the firm, since $U^N(\mathbf{p^N}) = 0$, becomes:

$$\max_{p_1,p_2} \Pi = S^N(\mathbf{p^N}) - U^N(\mathbf{p^N}) - (\tilde{U}(\mathbf{p^N}) - U^N(\mathbf{p^N})) = S^N(\mathbf{p^N}) - \Delta$$

Calculating the marginal prices that maximize the above expression we get Proposition 1.

**Proposition 1.** *<u>Monopoly:</u> If the consumer is naive habit forming the optimal marginal pricing scheme is:*

$$c = 0: \quad p_1^N = 0, \quad p_2^N > c$$

$$c > 0: \quad p_1^N < c, \quad p_2^N > c$$

*and the fixed fee $F^N$ equals to the perceived consumer surplus.*

**Proof**: See Appendix A

A firm facing a naive habit forming consumer has an incentive to distort the efficient allocation in order to maximize its profits. As the consumer misperceive her expected utility, the participation constraint is biased. As a result, the firm cannot maximize its profits by maximizing the social surplus and charging a fixed fee equal to the consumer surplus. This is because the perceived surplus is smaller than the surplus actually produced and thus the firm cannot extract the latter by means of a fixed fee. Therefore, the firm needs to distort the marginal prices by choosing the ones that maximize $S^N - \Delta$ and not just $S^N$. Thus, the undervaluation of the contract by the consumer explains why the firm charges prices different than the marginal cost.

The exact way in which the marginal prices are distorted depends on two characteristics of the consumer's behavior. Firstly, she is forward looking, and secondly, she underestimates the probability of consuming the second unit. As the consumer underestimates the probability of consuming the second unit the firm has an incentive to charge a price greater than the marginal cost, already shown by the literature on hyperbolic discounting and myopia. On the other hand, given that the consumer is forward looking and takes into consideration the opportunity cost of consuming the first unit, the firm finds it optimal to decrease below cost the marginal price of the first unit in order to constrict the downward bias in consumption. But more importantly, this marginal prices exacerbate the mistake of the second unit consumption. A distorted price below marginal cost makes more probable the consumption of the first unit. Moreover, it increases the probability of having a second consumption opportunity which means that the probability of consuming two units of the good increases as well. Thus, the consumption of the second unit becomes more probable not only because the consumer acquires a habit, she does not expect, but also because the first unit marginal price facilitates it.

For these reasons, the optimal pricing scheme when the consumer is naive habit forming resembles to the scheme we observe in several markets, namely *three part tariff.* This consist of a fixed fee, an included allowance of units for which the marginal price equals to zero and a positive marginal price for units beyond the allowance. When the marginal cost is equal to zero, the marginal price of the first unit is equal to zero and the marginal price of the second unit is greater than the marginal cost .

Interestingly, even if it seems that for the first unit there would be overconsumption, the fact that the consumer is forward looking and naive of her habit forming behavior produces the inverse result. For example, when the marginal cost is zero, $c = 0$, even if the marginal price of the first unit is zero its optimal threshold is positive thus there is underconsumption with respect to the efficient allocation of the sophisticated habit forming consumer. Moreover, the larger is the habit formation coefficient $\beta$, the greater is the first unit threshold $v_{1N}^*$ because the difference between the first and second unit optimal marginal price is larger.

The reason why there is underconsumption in the first period is because in this way the firm mitigates the undervaluation of the contract and thus decreases $\Delta$. In the profit function of the firm, there are two opposites effects. On the one hand, the firm wants to decrease the price below marginal cost in order to inflate the second period mistake, namely a price that would make more probable the consumption of the first unit and consequently of the second unit, which it could take advantage of. On the other hand, the firm has an incentive to increase the first period optimal threshold in order to minimize $\Delta$, that enters negatively into its profit function, and maximize its profits. Even if the firm can overcharge the second unit because of the mistaken beliefs, it cannot extract ex ante all the consumer surplus from this period. Therefore, the firm chooses a pricing scheme that balances the two opposite effects, and maximizes its profits.

Similarly, there is underconsumption of the second unit. The optimal second unit threshold for the naive consumer is always greater than the one of the sophisticated consumer, $v_2^{N*} > v_2^{S*}$. Thus, the consumer consumes less often than the efficient allocation of the sophisticated habit forming consumer.

As has already been mentioned, even if the consumer always consumes less than the optimal, she is left with a positive consumer surplus, because the firm cannot extract it all. This would give an incentive to the consumer to remain naive and not pay the cost of getting sophisticated and learning her true type. Remaining naive is beneficial for her both because the firm cannot extract all her surplus and because she avoids paying any information cost in order to become sophisticated.

A common critic is that naivety goes away with learning. Learning can mitigate naivety when appropriate feedback is provided (Bolger and Önkal-Atay, 2004). However, the firm may have no incentive to educate consumer and provide feedback Gabaix and Laibson (2006). Moreover, the consumers may learn slowly (Grubb and Osborne, 2015) or their lessons may be forgotten (Agarwal et al., 2013). Furthermore, three part tariffs are optimal even when consumers are partially naive (See Appendix B) which could resemble the period that she learns her true type.

**Oligopoly-Hotelling Competition:** In this subsection we introduce competition in order to examine if the pricing structure that is optimal under monopoly would be also optimal in an oligopolistic environment.

Let us consider a market with a continuum of naive habit forming consumers uniformly distributed on a Hotelling line and two firms $i = \{A, B\}$, positioned at the end points of the line.

The maximization problem of the firm $i$ is:

$$\max_{U_i} \Pi_i = D(U_i, U_{-i})(S_i^N(\mathbf{p_i^N}) - U_i^N(\mathbf{p_i^N}) + \Delta_i)$$

$$\text{s.t.} \quad U_i^N(\mathbf{p_i^N}) \geq 0,$$

where $D(U_i, U_{-i}) = \dfrac{U_i - U_{-i} + \tau}{2\tau}$ is firm i's market share [8]. The competition is in the utility space and $\tau$ is the transportation cost. Moreover, $S_i^N$ is the social surplus, $U_i^N$ is the mistakenly expected consumers surplus and $\Delta_i$ the difference between the actual and the mistakenly expected consumer surplus that are created by firm $i$, defined as in the case of the monopolist.

We know from Armstrong and Vickers (2001) that if there is strict full market coverage when firms set marginal prices optimally and charge markup $\tau$, then this is the equilibrium.

In this case, there is strict full market coverage when:

$$\frac{2}{3}(S_i + \Delta_i) \geq \tau$$

If we make the assumption that the above inequality holds and thus there is full market coverage in this market, then we have Proposition 2.

**Proposition 2.** *<u>Hotelling Duopoly</u>: Let $\tau$ be sufficiently small for strict full market coverage and the consumer be naive habit forming. Then, the optimal pricing scheme is:*

$$c = 0: \quad p_1^{N*} = 0, \quad p_2^{N*} > c, \quad F_i^{N*} = \frac{\tau}{2}$$

$$c > 0: \quad p_1^{N*} < c, \quad p_2^{N*} > c, \quad F_i^{N*} = \frac{\tau}{2}$$

[8] The consumer who is indifferent between the two firms is given by $U_i - \tau x = U_{-i} - \tau(1 - x) \Rightarrow x = \frac{U_i - U_{-i} + \tau}{2\tau}$

The marginal prices equal the monopolistic one. The competition among the firms is in utility space and thus it affects only the fixed fee charged as compared to the contract offered by the monopolist. The pricing scheme is exactly the same in both cases because there is full market coverage. Interestingly, a three part tariff contract is still the optimal pricing scheme when there is competition.

As one would expect, the more intense the competition, namely the smaller is $\tau$, the less scope there is for price discrimination.

## 4 Asymmetric Information

Now suppose that the firm cannot observe the type of the consumer. However, it is common knowledge that the probability that the consumer is sophisticated is $\gamma$.

We use the taxation principle, thus the screening is done with respect to the pricing scheme.[9] The firm offers a menu of contracts. Without any loss of generality, we can restrict the analysis to the case in which it offers two contracts as there are only two types. Let $\mathbf{p^N} = \{F^N, p_1^N, p_2^N\}$ and $\mathbf{p^S} = \{F^S, p_1^S, p_2^S\}$ be the contracts intended for the naive and the sophisticated consumer, respectively. This menu of tariffs completely identifies the allocation.

[9] The use of the taxation principle is more realistic and closer to what we observe. Moreover, the nature of the direct problem with multi-dimensional uncertainty makes the problem not tractable. The uncertainty is multi-dimensional because it concerns both the type of consumer at the contracting period, and the valuation of the good.

The maximization problem of the firm is:

$$\max_{\mathbf{p^S},\mathbf{p^N}} \quad \gamma(S^S(\mathbf{p^S}) - U^S(\mathbf{p^S})) + (1-\gamma)(S^N(\mathbf{p^N}) - U^N(\mathbf{p^N}) + \Delta)$$

$$\text{s.t.} \quad \begin{array}{ll} U^N(\mathbf{p^N}) \geq 0 & IR_N \\ U^S(\mathbf{p^S}) \geq 0 & IR_S \\ U^N(\mathbf{p^N}) \geq U^N(\mathbf{p^S}) & IC_N \\ U^S(\mathbf{p^S}) \geq U^S(\mathbf{p^N}) & IC_S \end{array}$$

$U^N(\mathbf{p^N}) \geq 0$ and $U^S(\mathbf{p^S}) \geq 0$ are the *participation constraints* of the naive and sophisticated consumer, respectively. Moreover, $U^N(\mathbf{p^N}) \geq U^N(\mathbf{p^S})$ and $U^S(\mathbf{p^S}) \geq U^S(\mathbf{p^N})$ are the *incentive compatibility constraints*: they say that each type should not have any incentive to mimic the other. Note that the participation constraint must hold ex ante. Once the consumer has signed the contract, she is obliged to comply for the whole contract period, even if she would have an incentive to deviate.

As we saw before, in the case of the full information, i.e. when the type is observable, the profit of the firm is greater when in the market there is only a sophisticated consumer. The profit in the case of the sophisticated consumer is the first best, because there is marginal cost pricing, first best allocation and with the fixed fee all the consumer surplus which in this case equals the social surplus is extracted. In the case of the naive habit forming consumer the firm finds it optimal to distort the allocation, as it cannot extract all the consumer surplus, charging $p_1 < c$, $p_2 > c$ and a fixed fee equal to the consumer's ex ante expected utility as is perceived at period zero, which is smaller than the consumer surplus of the sophisticated one.[10]

The above discussion suggests that the incentive constraint that we expect to bind at

[10] The relative ranking of optimal profit is important because it determines which market segment the firm would like to offer a discounted markup to.

the optimum is the one of the naive consumer. This is because the naive consumer at the contract period does not know that she will acquire a habit and that her utility will be greater than the one she expects to be. Marginal cost pricing creates a larger expected utility for the sophisticated consumer than for the naive consumer, thus the firm charges a fixed fee that the naive consumer would not be willing to pay.

On the other hand, the optimal full information contract is not incentive compatible for the sophisticated consumer, because she would prefer the contract of the naive consumer rather than her own first best allocation. Even if the marginal pricing is distorted, it allows her to enjoy a strictly positive surplus equal to $U^S(\mathbf{p^N}) - U^N(\mathbf{p^N})$. This suggests intuitively that it is the incentive constraint $IC_S$ that will be binding in the second-best problem. This intuition is confirmed formally in the following Lemma, which characterizes which constraints bind and which ones do not:

**Lemma 3.** *At the solution to the asymmetric information model, constraints $IR_N$ and $IC_S$ bind, whereas constraints $IR_S$ and $IC_N$ are redundant. More specifically:*

$$\begin{array}{ll} U^N(\mathbf{p^N}) = 0 & IR_N \\ U^S(\mathbf{p^S}) > 0 & IR_S \\ U^N(\mathbf{p^N}) > U^N(\mathbf{p^S}) & IC_N \\ U^S(\mathbf{p^S}) = U^S(\mathbf{p^N}) & IC_S \end{array}$$

**Proof**.

*$IR_S$ slack*: We show that if $IR_N$ and $IC_S$ hold at the optimum then $IR_S$ can be discarded.

$$U^S(\mathbf{p^S}) \geq U^S(\mathbf{p^N}) \geq U^N(\mathbf{p^N}) \geq 0 \Rightarrow U^S(\mathbf{p^S}) > 0$$

*$IR_N$ binds*: Otherwise increasing the fixed fee both of the sophisticated and the naive

consumer by a small positive $\epsilon$ would preserve the $IR_N$, would not affect the $IC_S$ and $IC_N$, and raise profits which a contradiction to $\mathbf{p^S}$ and $\mathbf{p^N}$ be optimal.

*$IC_S$ bind*: Suppose not, so that $U^S(\mathbf{p^S}) > U^S(\mathbf{p^N})$. Then the marginal prices for both the sophisticated and the naive consumer would be the optimal prices of the full information model because neither $IC_N$ nor $IC_S$ would bind. The firm could then increase the fixed fee of the sophisticated consumer without violating the $IC_S$, and increasing its profits. Thus, we expect that it binds at the optimum.

*$IC_N$ slacks*: Suppose not, so that $U^N(\mathbf{p^N}) = U^N(\mathbf{p^S})$ and let $\{\mathbf{p^{N\prime}}, \mathbf{p^{S\prime}}\}$ be a solution to the relaxed problem subject only to $IR_N$ and $IC_S$. Thus, if $IC_N$ does not slack then this solution violates it, namely the naive consumer prefers $\mathbf{p^{S\prime}}$ to $\mathbf{p^{N\prime}}$. Then, if the profits from the segment of naive consumers is greater than the one of sophisticated consumers, the firm can raise profits by giving both types a contract $\{p_1^S, p_2^S, F^S + \epsilon\}$, as $IR_S$ slack. If the opposite holds, the firm is better off by giving both types $\{p_1^N, p_2^N, F^N\}$. Thus, we have a contradiction and $IC_N$ can be discarded. ■

The fact that we expect the incentive compatibility constraint of the naive consumer to be satisfied (i.e. to be actually slack at the optimum) implies that there will be marginal cost pricing and first best allocation for the sophisticated consumer. If this was not true then setting $\{p_1^S, p_2^S\}$ equal to $\{c, c\}$ whereas keeping $U^S$ constant would keep the incentive compatibility and the participation constraint of the sophisticated unaffected and it would not violate the incentive constraint of the naive because it is relaxed. But this would increase the surplus and the profits of the firm, a contradiction.

Moreover, as we expect the incentive compatibility constraint of the sophisticated consumer to bind at the optimum, it could be written as:

$$U^S(\mathbf{p^S}) = U^S(\mathbf{p^N}) \Rightarrow U^S(\mathbf{p^S}) = U^S(\mathbf{p^N}) - U^N(\mathbf{p^N}) + U^N$$

Thus, taking into consideration Lemma 3 the relaxed problem is:

$$\max_{\mathbf{p^N}} \quad \Pi = -\gamma \underbrace{\left(U^S(\mathbf{p^N}) - U^N(\mathbf{p^N})\right)}_{\substack{\text{Information rent of} \\ \text{Sophisticated}}} + (1-\gamma)\left(S^N(\mathbf{p^N}) + \underbrace{U^N(\mathbf{p^N}) - \tilde{U}(\mathbf{p^N})}_{\substack{\text{Mis-perception rent} \\ \text{of Naive}}}\right)$$

The first part is the information rent left to sophisticated consumers and the second part is the profit made from naive consumers. Interestingly, both types of consumers are left with a rent and the firm cannot extract all their surplus. The sophisticated consumer has an information rent due to the asymmetry of information. The naive consumer, even if she has no incentive to deviate, is left with a *mis-perception rent*. This rent is due to the consumer's naivety, i.e. the fact that she does not know what her true level of utility will be if she consumes the first unit. Thus, she would not sign a more expensive contract at the contracting stage, and so she is left ex post with a *mis-perception rent* $\Delta$ that is bigger than her expected surplus at the contract period, $\Delta > U^N(\mathbf{p^N}) = 0$.

The solution of the relaxed maximization problem of the firm is described by Proposition 3

**Proposition 3.** *The optimal contract that the firm offers in order to screen between sophisticated and naive habit forming consumers is:*

- ***Sophisticated consumer:*** $p_1^S = c, \quad p_2^S = c, \quad F^S = U^S(\mathbf{p^S}) - U^S(\mathbf{p^N})$
- ***Naive Consumer:*** *if* $c = 0$ *then* $\{p_1^N = 0, \quad p_2^N > c\}$ *and if* $c > 0$ *then* $\{p_1^N < c, \quad p_2^N > c\}$ *when* $\beta$ *is relatively small, and the fixed fee,* $F^N$*, equals to the perceived consumer surplus of the naive consumer.*

**Proof**: See Appendix A

The firm offers a menu of contract consisting of a two part tariff for the sophisticated consumer and a three part tariff for the naive consumer. Qualitatively, the pricing patterns that are optimal under full information are still optimal under asymmetric information. If the fraction of sophisticated consumers is relatively small, then the firm finds it optimal to offer only the contract intended for naive consumers and vice versa.

It remains to check that all the constraints are met, and in particular that the incentive compatibility constraint of the naive consumers slack at the optimum. This is shown in the Appendix A.

# 5 Assuming Uniform Distribution

## Comparing Asymmetric Information vs Full Information Case

Assuming that the distribution of the valuation of the service is uniform allows us to make clear comparisons between the results of the full information case and the asymmetric information case[11].

The marginal prices of the contract of the sophisticated consumer $\{p_1^S, p_2^S\}$ remain equal to the marginal cost, whereas the fixed fee, $F^S$, decreases. Thus, the sophisticated consumer is better off in the presence of naive consumers. In this case, naive consumers exert a positive externality on the sophisticated consumers. On the other hand, the marginal prices for the naive consumer, $\{p_1^N, p_2^N\}$, are distorted upwards and the fixed fee, $F^N$ is lower. Thus, there are two opposing effects on the welfare of naive consumers. However, it can be shown that overall this type of consumer is worse off in the presence of the sophisticated ones. This is

[11] See Appendix for detailed calculations

shown in Figure 2[12]. This is due to the fact that the firm must make the contract intended for naive consumers not attractive for the sophisticated ones. Moreover, the misperception rent in the case of full information is bigger that the one in the asymmetric information case. The decrease of the expected utility of the naive consumer, $U_N$, is less than her true expected utility, $\tilde{U}$ thus the distortion of the allocation has a bigger effect on the true one. More specifically, we see that the derivatives with respect to the marginal prices are:

$$\frac{d\Delta}{dp_1} < 0 \quad \text{and} \quad \frac{d\Delta}{dp_2} < 0$$

This means that an increase in marginal prices decreases the mis-perception rent. The marginal prices are greater that in the full information case, and thus the naive consumer is worse off.

Moreover, as we see from Figure 2, the more habit forming are the naive consumers, the more they are harmed by the presence of sophisticated consumers. Intuitively, the rent left to the naive consumers when they are more strongly habit forming is greater, thus the contract must be distorted more heavily to make it less attractive for sophisticated consumers.

As discussed before, the naive consumer is less likely to consume in the first period than the sophisticated because she mistakenly believes that she is not habit forming. Comparing the three part tariff with marginal cost pricing, we see that the contract offered to the naive consumer exacerbates this mistake. The decrease of the first unit marginal price is not enough to correct the mistake, due to the increase in the second period marginal price and the fact that the consumer is forward looking. This means that the optimal first period threshold given $\mathbf{p^N}$ is greater than the one associated with marginal cost pricing:

[12] For the graphical representation, we have assumed that $c = 0.10$, $\gamma = 0.10$ in all the Figures of this section. However, the results hold more generally.

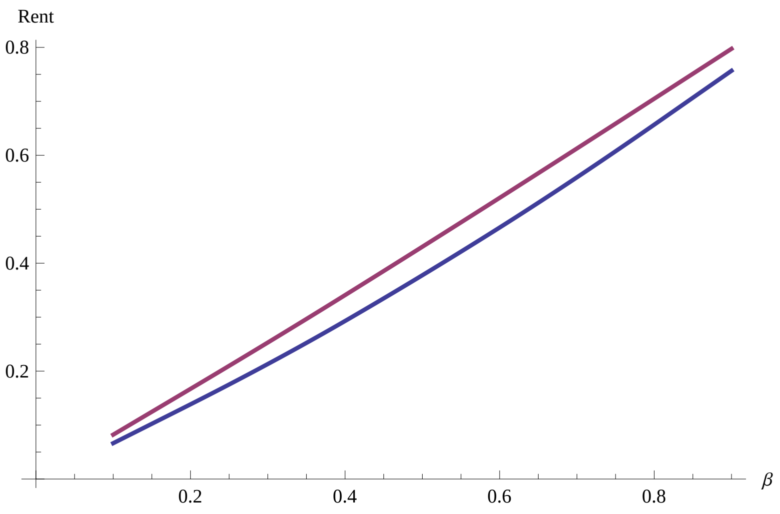



Figure 2: Rent of Naive Full Information > Rent of Naive Screening

$v^*_{1N}(\mathbf{p^N}) > v^*_{1N}(\mathbf{p^S}) = v^*_{1N}(\mathbf{c})$. This is shown in Figure 3. As expected, the sophisticated consumer is more likely to consume when she chooses the contract tailored for the naive consumer, $v^*_{1N}(\mathbf{p^S}) > v^*_{1S}(\mathbf{p^N})$, and even more likely when she chooses the contract tailored for her, $v^*_{1S}(\mathbf{p^N}) > v^*_{1S}(\mathbf{p^S})$. Moreover, as we see at Figure (3), the more habit forming is the consumer, the greater the exacerbation of the mistake.

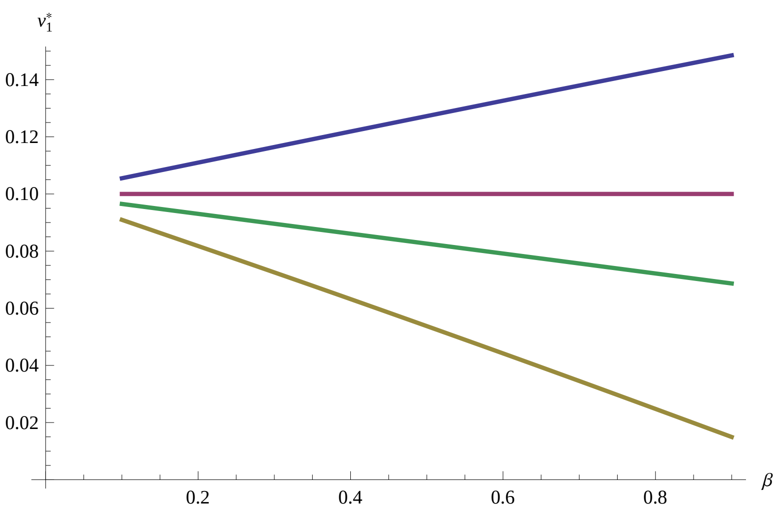



Figure 3: $v^*_{1N}(\mathbf{p^N}) > v^*_{1N}(\mathbf{p^S}) > v^*_{1S}(\mathbf{p^N}) > v^*_{1S}(\mathbf{p^S})$

Comparing now the optimal first-period threshold in the case of the full information,

$v^*_{1NF}(\mathbf{p_1^{NF}}, \mathbf{p_2^{NF}})$, with the one of the asymmetric information case, $v^*_{1NS}(\mathbf{p_1^{NS}}, \mathbf{p_2^{NS}})$, we see that there is an increase and thus there is more under-consumption. The firm distorts the marginal pricing for the naive consumer in order to make it less attractive for the sophisticated one. In this way, both marginal prices are increased, so the mistake gets larger (Figure 4). Moreover, the more habit forming is the consumer, i.e. the greater the habit forming coefficient $\beta$, the bigger the exacerbation of the mistake.

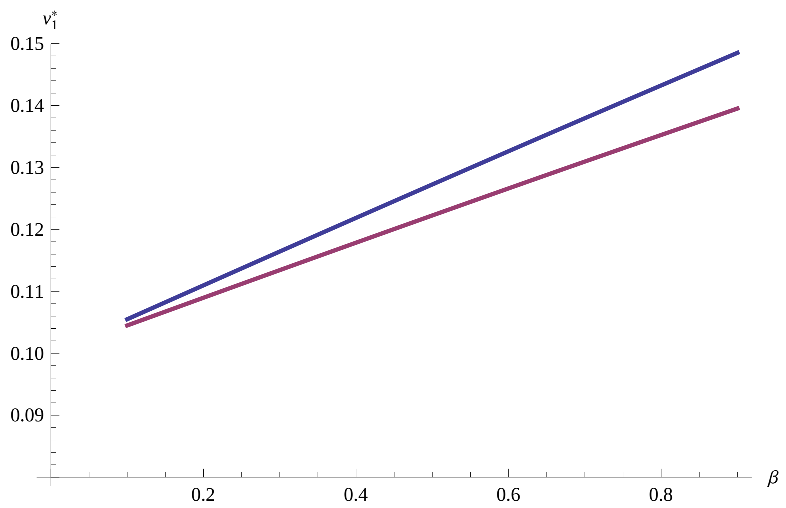


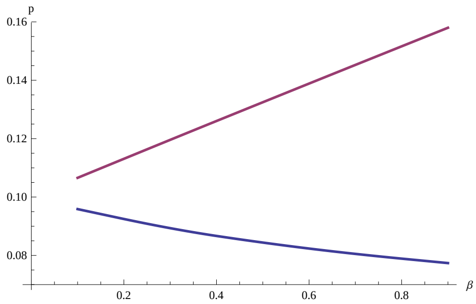

Figure 4: $v^*_{1NS}(\mathbf{p_1^{NS}}, \mathbf{p_2^{NS}}) > v^*_{1NF}(\mathbf{p_1^{NF}}, \mathbf{p_2^{NF}})$

The profits of the firms decrease with respect to the full information case both for the sophisticated and the naive consumer. The fact that the firm cannot exploit the naivety of the consumer and at the same time cannot observe her type, decreases its profits.

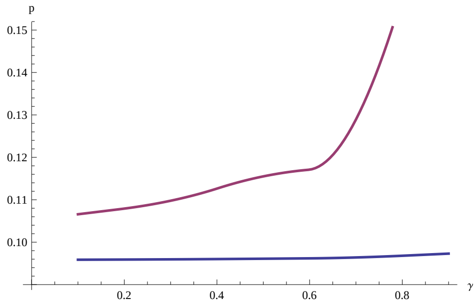
### Comparative Statics

Let now consider how the optimal contracts would change when the parameters of the model change. The marginal prices of the sophisticated consumer would remain unchanged, equal to the marginal cost. What changes is the fixed fee for sophisticated consumers, $F^S$, and the whole contract of the naive ones, $\mathbf{p^N}$.

The more habit forming consumers are, i.e. the greater habit forming coefficient $\beta$, the lower is the first unit price, $p_1^N$, and the greater is the second unit price, $p_2^N$ (Figure 5a). Thus, the more habit forming are consumers, the greater is the difference between the marginal prices of the two units.




(a) $\frac{dp_{1N}}{d\beta} < 0 \quad \frac{dp_{2N}}{d\beta} > 0$




(b) $\frac{dp_{1N}}{d\gamma} > 0 \quad \frac{dp_{2N}}{d\gamma} > 0$

Figure 5: Comparative Statics

Finally, the more sophisticated consumers there are in the market, i.e. the greater is $\gamma$, the greater are both the first unit marginal price, $p_1^N$, and the second unit marginal price, $p_2^N$ (Figure 5b).

## 6 Conclusion

The last decades, the provision of menu of contracts consisting of two part tariff and three part tariff has become prevalent in a number of markets. Moreover, habit formation, and there is evidence that the consumption of communication services like cell phones and internet are habit forming. Furthermore, in these kind of markets, “three part tariff” contracts, which comprise a fixed fee, an allowance of free units, and a positive price for additional units beyond the allowance, are becoming increasingly popular.

This article claims that habit forming behavior is an important characteristic of this kind of markets and can play a significant role in explaining the pricing schemes that are observed. In particular, naive habit formation makes it optimal for the firm to charge "three part tariff". We show that if the consumption choice is made sequentially within the contract period, the consumer undervaluates the offered contract at the contracting period and underestimates the demand conditional on it being high, which are the main characteristics of a naive habit forming behavior, it is optimal for the firm to offer a three part tariff.

This explanation could be viewed as an alternative channel to the overconfidence model of Grubb (2009) that also explains this kind of pricing scheme. We show that if the elements mentioned before are present, it is enough that the consumer underestimates the demand conditional on it being high for the introduction of three part tariff to be optimal.

Interestingly, the firm cannot take advantage of the naivety of its client. To the contrary, it is worse off when it encounters a naive habit forming consumer with respect to a sophisticated one. This is because the firm cannot charge a fixed fee that extracts all the consumer surplus at the contracting stage. If possible, the firm would have an incentive to inform the consumer about her naivety but this would make the consumer worst off. On the other hand, the consumer has no incentive to pay the cost of getting informed about her own type.

Interestingly, even if the optimal marginal price for low volumes is smaller than the marginal cost, or even zero, a naive consumer underconsumes compared to a sophisticated one. The first period optimal threshold $v_1^*$ is a function of both units' marginal prices. The firm finds it optimal to charge a second unit price above marginal cost, even if this would decrease first period consumption because the second period is when the firm could take advantage of the mistaken expectations of the consumer. At the same time, however, it is the period when the consumer cannot foresee the real value of the good, and thus the period when the firm cannot extract the whole consumer surplus ex ante. The firm through its pricing scheme $\{p_1, p_2, F\}$ wants to have a balance between its incentive to take advantage of the consumer's mistake and to minimize $\Delta$, namely the part of consumer surplus that cannot extract. Thus, in order to mitigate the undervaluation of the whole contract and maximize its profits it does not charge a first unit price $p_1$ that would restore the demand at the level that a sophisticated consumer would have. Contrary, the firm finds it optimal to charge prices that boost demand in the first period but not as much as to overconsume or achieve efficiency for low volumes. Similarly, for high volumes the consumer underconsumes. Thus, the firm by charging this kind of contract exacerbates the mistake the consumer does because of her naivety in order to mitigate the profit losses it has.

Moreover, this article claims that the obseved menu of contracts could be explained by the existence of habit forming consumers of diverse sophistication about their habit forming behavior. It is shown that the firm finds it optimal to offer two part tariff to sophisticated consumers and three part tariff to naive ones.

The presence of naive consumers in the market exerts a positive externality to the sophisticated consumers instead of the other way around. In the full information case the firm can extract all the consumer surplus of the sophisticated consumer. In the asymmetric information case, though, they are left with the information rent, as a result of their superior information and the objective of the firm to screen between the two types.

This come as a result of an interesting implication of this model which is that the sophisticated consumer has an incentive to pretend the naive consumer and choose the contract intended for them. It seems counter intuitive because this contract charges highly the high consumption and the sophisticated consumer knows that she will form a habit that will make high consumption more likely. Thus, we see that she finds optimal to mimic the naive, and consume less because she is left with a rent. For this reason, the firm finds it optimal to leave information rent to the sophisticated consumers and make the contract intended for them incentive compatible.

However, both types are left with a rent. The naive consumers are left with a misperception rent and the sophisticated an information rent. The naive consumers are ex post worst off in the presence of sophisticated consumers, since the objective of the firm to make the contract intended for naive consumers less attractive to sophisticated ones leads to an increase of the marginal prices and thus a decrease to the ex post misperception rent.

## References


Abel, A. B. (1990). Asset prices under habit formation and catching up with the joneses. *The American Economic Review*, 38–42.

Agarwal, S., J. C. Driscoll, X. Gabaix, and D. Laibson (2013). Learning in the credit card market. *Available at SSRN 1091623*.

Armstrong, M. and J. Vickers (2001). Competitive price discrimination. *RAND Journal of Economics*, 579–605.

Armstrong, M. and J. Vickers (2012). Consumer protection and contingent charges. *Journal of Economic Literature 50*(2), 477–493.

Becker, G., M. Grossman, and K. Murphy (1991). Rational addiction and the effect of price on consumption. *The American Economic Review 81*(2), 237–241.

Becker, G. and K. Murphy (1988). A theory of rational addiction. *The Journal of Political Economy*, 675–700.

Bianchi, A. and J. G. Phillips (2005). Psychological predictors of problem mobile phone use. *CyberPsychology & Behavior 8*(1), 39–51.

Boldrin, M., L. J. Christiano, and J. D. Fisher (2001). Habit persistence, asset returns, and the business cycle. *American Economic Review*, 149–166.

Bolger, F. and D. Önkal-Atay (2004). The effects of feedback on judgmental interval predictions. *international Journal of forecasting 20*(1), 29–39.

Campbell, J. Y. and J. H. Cochrane (1999). Vby force of habit: A consumption based explanation of aggregate stock market behaviorv. *Journal of Political Economy 107*(2), 205–251.

Carbajal, J. C. and J. Ely (2012). Optimal contracts for loss averse consumers. Technical report, University of Queensland, School of Economics.

Carroll, C. D., J. Overland, and D. N. Weil (2000). Saving and growth with habit formation. *American Economic Review*, 341–355.

Constantinides, G. M. (1990). Habit formation: A resolution of the equity premium puzzle. *Journal of political Economy*, 519–543.

DellaVigna, S. and U. Malmendier (2004). Contract design and self-control: Theory and evidence. *The Quarterly Journal of Economics 119*(2), 353.

Driskill, R. and S. McCafferty (2001). Monopoly and oligopoly provision of addictive goods. *International Economic Review 42*(1), 43–72.

Eliaz, K. and R. Spiegler (2006). Contracting with diversely naive agents. *The Review of Economic Studies 73*(3), 689–714.

Eliaz, K. and R. Spiegler (2008). Consumer optimism and price discrimination. *Theoretical Economics 3*(4), 459–497.

Esteban, S., E. Miyagawa, and M. Shum (2007). Nonlinear pricing with self-control preferences. *Journal of Economic Theory 135*(1), 306–338.

Fethke, G. and R. Jagannathan (1996). Habit persistence, heterogeneous tastes, and imperfect competition. *Journal of Economic Dynamics and Control 20*(6-7), 1193–1207.

Fuhrer, J. C. (2000). Habit formation in consumption and its implications for monetary-policy models. *American Economic Review*, 367–390.

Gabaix, X. and D. Laibson (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics 121*(2), 505–540.

Grubb, M. (2009). Selling to overconfident consumers. *The American Economic Review 99*(5), 1770–1807.

Grubb, M. D. (2014). Consumer inattention and bill-shock regulation. *The Review of Economic Studies*, rdu024.

Grubb, M. D. and M. Osborne (2015). Cellular service demand: Biased beliefs, learning, and bill shock. *American Economic Review 105*(1), 234–271.

Hahn, J.-H., J. Kim, S.-H. Kim, and J. Lee (2012). Screening loss averse consumers. Technical report, mimeo.

Heidhues, P. and B. Kőszegi (2010). Exploiting naivete about self-control in the credit market. *The American Economic Review*, 2279–2303.

Herweg, F. and K. Mierendorff (2013). Uncertain demand, consumer loss aversion, and flat rate tariffs. *Journal of the European Economic Association 11*(2), 399–432.

Hooper, V. and Y. Zhou (2007). Addictive, dependent, compulsive? a study of mobile phone usage. In *Conference Proceedings of the 20th Bled eConference eMergence held in Bled, Slovenia: Merging and Emerging Technologies, Processes, and Institutions.*

Jermann, U. J. (1998). Asset pricing in production economies. *Journal of Monetary Economics 41*(2), 257–275.

Lambrecht, A. and B. Skiera (2006). Paying too much and being happy about it: Existence, causes, and consequences of tariff-choice biases. *Journal of marketing Research 43*(2), 212–223.

Loewenstein, G., T. O'Donoghue, and M. Rabin (2003). Projection bias in predicting future utility. *The Quarterly Journal of Economics 118*(4), 1209.

Maskin, E. and J. Riley (1984). Monopoly with incomplete information. *The RAND Journal of Economics 15*(2), 171–196.

Miao, C.-H. (2010). Consumer myopia, standardization and aftermarket monopolization. *European economic review 54*(7), 931–946.

Muellbauer, J. (1988). Habits, rationality and myopia in the life cycle consumption function. *Annales d'Economie et de Statistique*, 47–70.

Mussa, M. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic theory 18*(2), 301–317.

Nakamura, E. and J. Steinsson (2011). Price setting in forward-looking customer markets. *Journal of Monetary Economics 58*(3), 220–233.

Park, W. (2005). Mobile phone addiction. *Mobile Communications*, 253–272.

Piccione, M. and A. Rubinstein (2003). Modeling the economic interaction of agents with diverse abilities to recognize equilibrium patterns. *Journal of the European economic association 1*(1), 212–223.

Pollak, R. (1970). Habit formation and dynamic demand functions. *The Journal of Political Economy 78*(4), 745–763.

Rubinstein, A. (1993). On price recognition and computational complexity in a monopolistic model. *Journal of Political Economy*, 473–484.

Sandroni, A. and F. Squintani (2010). Overconfidence and adverse selection: the case of insurance. Technical report, working paper.

Spinnewijn, J. (2013). Insurance and perceptions: how to screen optimists and pessimists. *The Economic Journal 123*(569), 606–633.

# 7 Appendix A

**Optimal Pricing Benchmark model**

The optimization problem of the firm when the consumer is non habit forming is:

$$\max_{p_1,p_2} \Pi = S - U \quad \text{s.t.} \quad U = 0$$

and the optimal consumption rule is:

$$v_1^* = p_1 + \int_{p_1}^{p_2} (1 - F(v_2))dv_2$$

Then the first order conditions are:

$$\frac{\partial \Pi}{\partial p_1} = \frac{\partial S}{\partial p_1} + \frac{\partial S}{\partial v_1^*}\frac{\partial v_1^*}{\partial p_1} = 0$$
$$\frac{\partial \Pi}{\partial p_2} = \frac{\partial S}{\partial p_2} + \frac{\partial S}{\partial v_1^*}\frac{\partial v_1^*}{\partial p_2} = 0$$

and

$$\frac{\partial S}{\partial p_1} = (-p_1 + c)f(p_1)F(v_1^*)$$
$$\frac{\partial S}{\partial v_1^*} = f(v_1^*)\left(-v_1^* + c - \int_{p_2}^{1}(v_2 - c)f(v_2)dv_2 + \int_{p_1}^{1}(v_2 - c)f(v_2)dv_2\right)$$
$$\frac{\partial v_1^*}{\partial p_1} = 1 - (1 - F(p_1)) = F(p_1)$$
$$\frac{\partial S}{\partial p_2} = (-p_2 + c)f(p_2)(1 - F(v_1^*))$$
$$\frac{\partial v_1^*}{\partial p_2} = 1 - F(p_2)$$

Substituting the above partial derivatives into the first order conditions, and then solving the system of equations we prove that the optimal marginal prices are $\{p_1, p_2\} = \{c, c\}$

**Proof of Lemma 2**

The optimization problem of the firm when the consumer is non habit forming is:

$$\max_{p_1,p_2} \Pi^S = S^S - U^S \quad \text{s.t.} \quad U^S = 0$$

and the optimal consumption rule is:

$$v_{1S}^* = p_1 - \int_{p_2-\beta}^{p_1} (1 - F(v_2))dv_2$$

Then the first order conditions are:

$$\frac{\partial \Pi^S}{\partial p_1} = \frac{\partial S^S}{\partial p_1} + \frac{\partial S^S}{\partial v_{1S}^*} \frac{\partial v_{1S}^*}{\partial p_1} = 0$$
$$\frac{\partial \Pi^S}{\partial p_2} = \frac{\partial S^S}{\partial p_2} + \frac{\partial S^S}{\partial v_{1S}^*} \frac{\partial v_{1S}^*}{\partial p_2} = 0$$

and

$$\frac{\partial S^S}{\partial p_1} = (-p_1 + c)f(p_1)F(v_{1S}^*)$$
$$\frac{\partial S^S}{\partial v_{1S}^*} = f(v_1^*) \left( -v_1^* + c - \int_{p_2-\beta}^{1} (v_2 - c)f(v_2)dv_2 + \int_{p_1}^{1} (v_2 - c)f(v_2)dv_2 \right)$$
$$\frac{\partial v_{1S}^*}{\partial p_1} = F(p_1)$$
$$\frac{\partial S^S}{\partial p_2} = (-p_2 - \beta + c)f(p_2 - \beta)(1 - F(v_1^*))$$
$$\frac{\partial v_{1S}^*}{\partial p_2} = 1 - F(p_2 - \beta)$$

Substituting the above partial derivatives into the first order conditions, and then solving the system of equations we prove that the optimal marginal prices are $\{p_1, p_2\} = \{c, c\}$

**Proof of Proposition 1**

Thus the optimization problem of the firm is:

$$\max_{U^*, p_1, p_2} \Pi = S^S - U^N + (U^N - U^S) = S - U^N + \Delta \quad \text{s.t.} \quad U^N \geq 0$$

and optimal consumption rule is:

$$v_{1N}^* = p_1 + \int_{p_1}^{p_2} (1 - F(v_2))dv_2$$

The expected gross surplus is the one produced in a market with a habit forming consumer.

$$
\begin{aligned}
S &= \int_{v_{1N}^*}^{1} (v_1 - c)dF(v_1) + F(v_{1N}^*)\int_{p_1}^{1}(v_2 - c)dF(v_2) + \int_{v_{1N}^*}^{1}\int_{p_2-\beta}^{1}(v_2 + \beta - c)f(v_2)dv_2dF(v_1) \\
&= \int_{v_{1N}^*}^{1} (v_1 - c)dF(v_1) + F(v_{1N}^*)\int_{p_1}^{1}(v_2 - c)dF(v_2) + \int_{v_{1N}^*}^{1}\int_{p_2-\beta}^{1}(v_2 + \beta)f(v_2)dv_2dF(v_1) \\
&- c\int_{v_{1N}^*}^{1} f(v_1)(1 - F(p_2 - \beta))
\end{aligned}
$$

Moreover, $\Delta$ is the difference between the perceived and the optimal utility of the consumer.

$$
\Delta = U^N - U^S = (1 - F(v_{1N}^*))\int_{p_2}^{1}(1 - F(v_2))dv_2 + p_2\int_{v_{1N}^*}^{1} f(v_1)(1 - F(p_2 - \beta))dv_1
$$

Simplifying and deleting $\int_{v_{1N}^*}^{1}\int_{p_2-\beta}^{1}(v_2+\beta)f(v_2)dv_2dF(v_1)$ from $S$ and $\Delta$ then the first order conditions are:

with respect to $p_1$:

$$
\frac{\partial \Pi}{\partial p_1} = \frac{\partial S}{\partial v_{1N}^*}\frac{\partial v_{1N}^*}{\partial p_1} + \frac{\partial \Delta}{\partial v_{1N}^*}\frac{\partial v_{1N}^*}{\partial p_1} + \frac{\partial \Delta}{\partial p_1}
$$

$$
\begin{aligned}
\frac{\partial S}{\partial v_{1N}^*} &= \left(-v_{1N}^* + c + \int_{p_1}^{1}(v_2 - c)dF(v_2) + c(1 - F(p_2 - \beta))\right) f(v_{1N}^*) \\
\frac{\partial \Delta}{\partial v_{1N}^*} &= \left(-\int_{p_2}^{1}(1 - F(v_2))dv_2 - p_2(1 - F(p_2 - \beta))\right) f(v_{1N}^*) \\
\frac{\partial v_{1N}^*}{\partial p_1} &= 1 - (1 - F(p_1)) = F(p_1) \\
\frac{\partial S}{\partial p_1} &= -F(v_{1N}^*)(p_1 - c)f(p_1)
\end{aligned}
$$

Then the first order condition is:

$$\frac{\partial \Pi}{\partial p_1} = \left(-v_{1N}^* + c + \int_{p_1}^{1}(v_2 - c)dF(v_2) + c(1 - F(p_2 - \beta))\right)F(p_1) - \frac{F(v_{1N}^*)f(p_1)}{f(v_{1N}^*)}(p_1 - c)$$
$$- F(p_1)\int_{p_2}^{1}(1 - F(v_2))dv_2 - p_2(1 - F(p_2 - \beta))F(p_1) =$$
$$= -v_{1N}^* + c + \int_{p_1}^{1}(v_2 - c)dF(v_2) + c(1 - F(p_2 - \beta) - \frac{F(v_{1N}^*)f(p_1)}{f(v_{1N}^*)F(p_1)}(p_1 - c)$$
$$- \int_{p_2}^{1}(1 - F(v_2))dv_2 - p_2(1 - F(p_2 - \beta)) =$$
$$= -p_1 - \int_{p_1}^{p_2}(1 - F(v_2))dv_2 + c + \int_{p_1}^{1}(v_2 - c)dF(v_2) + c(1 - F(p_2 - \beta))$$
$$- \frac{F(v_{1N}^*)f(p_1)}{f(v_{1N}^*)F(p_1)}(p_1 - c) - \int_{p_2}^{1}(1 - F(v_2))dv_2 - p_2(1 - F(p_2 - \beta)) =$$
$$= p_1 + c - 1 + p_1 + 1 - p_1F(p_1) - c(1 - F(p_1)) + c(1 - F(p_2 - \beta))$$
$$- \frac{F(v_{1N}^*)f(p_1)}{f(v_{1N}^*)F(p_1)}(p_1 - c) - p_2(1 - F(p_2 - \beta))$$

Then

$$p_1\left(F(p_1) + \frac{F(v_{1N}^*)f(p_1)}{f(v_{1N}^*)F(p_1)}\right) = c\left(F(p_1) + 1 - F(p_2 - \beta) + \frac{F(v_{1N}^*)f(p_1)}{f(v_{1N}^*)F(p_1)}\right)$$
$$- p_2(1 - F(p_2 - \beta))$$

$$p_1\left(\frac{F(p_1)^2 f(v_{1N}^*) + F(v_{1N}^*)f(p_1)}{F(p_1)f(v_{1N}^*)}\right) = c\left(\frac{F(p_1)^2 f(v_N) + F(v_{1N}^*)f(p_1)}{F(p_1)f(v_{1N}^*)}\right) + c(1 - F(p_2 - \beta))$$
$$- p_2(1 - F(p_2 - \beta))$$

Thus

$$p_1 = c - (p_2 - c)\left(\frac{F(p_1)f(v_{1N}^*)(1 - F(p_2 - \beta))}{F(p_1)^2 f(v_{1N}^*) + F(v_{1N}^*)f(p_1)}\right) \tag{5}$$

with respect to $p_2$:

$$\frac{\partial \Pi}{\partial p_2} = \frac{\partial S}{\partial v_{1N}^*}\frac{\partial v_{1N}^*}{\partial p_2} + \frac{\partial \Delta}{\partial v_{1N}^*}\frac{\partial v_{1N}^*}{\partial p_2} + \frac{\partial \Delta}{\partial p_2}$$

$$\frac{\partial S}{\partial p_2} = c\int_{v_{1N}^*}^{1} f(v_1)f(p_2 - \beta)dv_1$$

$$\frac{\partial \Delta}{\partial p_2} = -(1 - F(v_{1N}^*))(1 - F(p_2)) + \int_{v_{1N}^*}^{1} f(v_1)(1 - F(p_2 - \beta))dv_1 - p_2\int_{v_{1N}^*}^{1} f(v_1)f(p_2 - \beta)dv_1$$

$$\frac{v_{1N}^*}{p_2} = 1 - F(p_2)$$

$$\frac{\partial \Pi}{\partial p_2} = \left(-v_{1N}^* + c + \int_{p_1}^{1} (v_2 - c) dF(v_2) + c(1 - F(p_2 - \beta))\right) f(v_{1N}^*)(1 - F(p_2))$$
$$+ c \int_{v_{1N}^*}^{1} f(v_1) f(p_2 - \beta) dv_1 - (1 - F(v_{1N}^*))(1 - F(p_2))$$
$$+ \left(- \int_{p_2}^{1} (1 - F(v_2)) dv_2 - p_2(1 - F(p_2 - \beta))\right) f(v_{1N}^*)(1 - F(p_2))$$
$$+ \int_{v_{1N}^*}^{1} f(v_1)(1 - F(p_2 - \beta v_1)) dv_1 - p_2 \int_{v_{1N}^*}^{1} f(v_1) f(p_2 - \beta) dv_1$$
$$= -v_{1N}^* + c + \int_{p_1}^{1} (v_2 - c) dF(v_2) + c(1 - F(p_2 - \beta))$$
$$+ c \left( \frac{\int_{v_{1N}^*}^{1} f(v_1) f(p_2 - \beta) dv_1}{f(v_{1N}^*)(1 - F(p_2))} \right) - \int_{p_2}^{1} (1 - F(v_2)) dv_2 - p_2(1 - F(p_2 - \beta))$$
$$- \frac{1 - F(v_{1N}^*)}{f(v_{1N}^*)} + \frac{\int_{v_{1N}^*}^{1} f(v_1)(1 - F(p_2 - \beta)) dv_1}{f(v_{1N}^*)(1 - F(p_2))} - p_2 \left( \frac{\int_{v_{1N}^*}^{1} f(v_1) f(p_2 - \beta) dv_1}{f(v_{1N}^*)(1 - F(p_2))} \right) = 0$$

Substituting for the optimal threshold and after some algebra [13]

$$\frac{\partial \Pi}{\partial p_2} = c - p_1 F(p_1) - c(1 - F(p_1)) + c(1 - F(p_2 - \beta)) + c \left( \frac{\int_{v_{1N}^*}^{1} f(v_1) f(p_2 - \beta) dv_1}{f(v_{1N}^*)(1 - F(p_2))} \right)$$
$$- p_2(1 - F(p_2 - \beta)) - \frac{1 - F(v_{1N}^*)}{f(v_{1N}^*)} + \frac{\int_{v_{1N}^*}^{1} f(v_1)(1 - F(p_2 - \beta)) dv_1}{f(v_{1N}^*)(1 - F(p_2))}$$
$$- p_2 \left( \frac{\int_{v_{1N}^*}^{1} f(v_1) f(p_2 - \beta) dv_1}{f(v_{1N}^*)(1 - F(p_2))} \right) = 0$$

[13] Let $p_1 < p_2$

$$- \int_{p_1}^{p_2} (1 - F(v_2)) dv_2 - \int_{p_2}^{1} (1 - F(v_2)) dv_2 - \int_{p_1}^{1} F(v_2) dv_2 = -1 + p_1$$
$$\int_{p_1}^{1} (v_2 - c) dF(v_2) = 1 - p_1 F(p_1) - \int_{p_1}^{1} F(v_2) dv_2$$

Then:

$$p_2\left(1-F(p_2-\beta)+\left(\frac{\int_{v_{1N}^*}^1 f(v_1)f(p_2-\beta)dv_1}{f(v_{1N}^*)(1-F(p_2))}\right)\right)=$$

$$-p_1F(p_1)+\frac{\int_{v_{1N}^*}^1 f(v_1)(1-F(p_2-\beta))dv_1-(1-F(v_{1N}^*))(1-F(p_2))}{f(v_{1N}^*)(1-F(p_2))}$$

$$+c\left(1-1+F(p_1)+1-F(p_2-\beta)+\left(\frac{\int_{v_{1N}^*}^1 f(v_1)f(p_2-\beta)dv_1}{f(v_{1N}^*)(1-F(p_2))}\right)\right)$$

Substituting (5) and rearranging:

$$p_2\left(1-F(p_2-\beta)+\left(\frac{\int_{v_{1N}^*}^1 f(v_1)f(p_2-\beta)dv_1}{f(v_{1N}^*)(1-F(p_2))}\right)-\frac{F(p_1)^2f(v_{1N}^*)(1-F(p_2-\beta))}{F(p_1)^2f(v_{1N}^*)+F(v_{1N}^*)f(p_1)}\right)=$$

$$+c\left(1-F(p_2-\beta)+\left(\frac{\int_{v_{1N}^*}^1 f(v_1)f(p_2-\beta))dv_1}{f(v_{1N}^*)(1-F(p_2))}\right)-\frac{F(p_1)^2f(v_{1N}^*)(1-F(p_2-\beta))}{F(p_1)^2f(v_{1N}^*)+F(v_{1N}^*)f(p_1)}\right)$$

$$+\frac{\int_{v_{1N}^*}^1 f(v_1)(1-F(p_2-\beta))dv_1-(1-F(v_{1N}^*))(1-F(p_2))}{f(v_{1N}^*)(1-F(p_2))}$$

Moreover, let for simplicity

$$A=1-F(p_2-\beta)+\left(\frac{\int_{v_{1N}^*}^1 f(v_1)f(p_2-\beta)dv_1}{f(v_{1N}^*)(1-F(p_2))}\right)-\frac{F(p_1)^2f(v_{1N}^*)(1-F(p_2-\beta))}{F(p_1)^2f(v_{1N}^*)+F(v_{1N}^*)f(p_1)}$$

$$=\left(\frac{\int_{v_{1N}^*}^1 f(v_1)f(p_2-\beta)dv_1}{f(v_{1N}^*)(1-F(p_2))}\right)+(1-F(p_2-\beta))\left(1-\frac{F(p_1)^2f(v_{1N}^*)}{F(p_1)^2f(v_{1N}^*)+F(v_{1N}^*)f(p_1)}\right)$$

$$=\left(\frac{\int_{v_{1N}^*}^1 f(v_1)f(p_2-\beta)dv_1}{f(v_{1N}^*)(1-F(p_2))}\right)+(1-F(p_2-\beta))\left(\frac{F(p_1)^2f(v_{1N}^*)+F(v_{1N}^*)f(p_1)}{F(p_1)^2f(v_{1N}^*)+F(v_{1N}^*)f(p_1)}\right)$$

$$+(1-F(p_2-\beta))\left(\frac{-F(p_1)^2f(v_{1N}^*)}{F(p_1)^2f(v_{1N}^*)+F(v_{1N}^*)f(p_1)}\right)$$

$$=\frac{\int_{v_{1N}^*}^1 f(v_1)f(p_2-\beta)dv_1}{f(v_{1N}^*)(1-F(p_2))}+\frac{(1-F(p_2-\beta))F(v_{1N}^*)f(p_1)}{F(p_1)^2f(v_{1N}^*)+F(v_{1N}^*)f(p_1)}>0$$

and

$$B=\frac{\int_{v_{1N}^*}^1 f(v_1)(1-F(p_2-\beta))dv_1-(1-F(v_{1N}^*))(1-F(p_2))}{(f(v_{1N}^*)(1-F(p_2)))}>0$$

Then the optimal price for the second quantity is:

$$p_2=c+\frac{B}{A} \tag{6}$$

Finally, substituting (6) back to (5), we get:

$$\begin{aligned}
p_1 &= c - (p_2 - c)\left(\frac{F(p_1)f(v_{1N}^*)(1-F(p_2-\beta))}{F(p_1)^2 f(v_{1N}^*) + F(v_{1N}^*)f(p_1)}\right)\\
&= c - (c + \frac{B}{A} - c)\left(\frac{F(p_1)f(v_{1N}^*)(1-F(p_2-\beta))}{F(p_1)^2 f(v_{1N}^*) + F(v_{1N}^*)f(p_1)}\right)\\
&= c - \frac{B}{A}\left(\frac{F(p_1)f(v_{1N}^*)(1-F(p_2-\beta))}{F(p_1)^2 f(v_{1N}^*) + F(v_{1N}^*)f(p_1)}\right)
\end{aligned}$$

Moreover if c=0 since $p_1$ cannot be negative since $F(p_1)$ cannot be negative then $p_1 = 0$

## Proof of Proposition 3

The profits of Naive are:

$$\begin{aligned}
\Pi^N &= S^N(\mathbf{p^N}) - U^N + \Delta\\
&= \int_{v_{1N}^*}^1 (v_1 - c)f(v_1)dv_1 + F(v_{1N}^*)\int_{p_1}^1 (v_2 - c)f(v_2)dv_2\\
&+ (p_2 - c)(1 - F(v_{1N}^*))(1 - F(p_2 - \beta)) + (1 - F(v_{1N}^*))\int_{p_2}^1 \big(1 - F(v_2))dv_2
\end{aligned}$$

where

$$\begin{aligned}
S^N(\mathbf{p^N}) &= \int_{v_{1N}^*}^1 \big(v_1 - c + \int_{p_{2N}-\beta}^1 (v_2 + \beta - c)dv_2\big)f(v_1)dv_1 + F(v_{1N}^*)\int_{p_{1N}}^1 (v_2 - c)f(v_2)dv_2 =\\
&= \int_{v_{1N}^*}^1 \big(v_1 - c)f(v_1)dv_1 + (1 - F(v_{1N}^*))\int_{p_{2N}-\beta}^1 (v_2 + \beta - c)dv_2\big)\\
&+ F(v_{1N}^*)\int_{p_{1N}}^1 (v_2 - c)f(v_2)dv_2
\end{aligned}$$

$$\begin{aligned}
U^S(\mathbf{p^N}) &= \int_{v_{1S}^*}^1 \big(v_1 - p_{1N} + \int_{p_{2N}-\beta}^1 (v_2 + \beta - p_{2N})dv_2\big)f(v_1)dv_1\\
&+ F(v_{1S}^*)\int_{p_{1N}}^1 (v_2 - p_{1N})f(v_2)dv_2 - F^N =\\
&= \int_{v_{1S}^*}^1 \big(v_1 - p_{1N})f(v_1)dv_1 + (1 - F(v_{1S}^*))\int_{p_{2N}-\beta}^1 (1 - F(v_2))dv_2\\
&+ F(v_{1S}^*)\int_{p_{1N}}^1 (1 - F(v_2))dv_2 - F^N
\end{aligned}$$

$$
\begin{aligned}
U^N(\mathbf{p^N})) = &\int_{v_{1N}^*}^1 \big(v_1 - p_{1N} + \int_{p_{2N}}^1 (v_2 - p_{2N}) f(v_2) dv_2\big) f(v_1) dv_1 \\
&+ F(v_{1N}^*) \int_{p_{1N}}^1 (v_2 - p_{1N}) f(v_2) dv_2 - F^N = \\
&= \int_{v_{1N}^*}^1 \big(v_1 - p_{1N}) f(v_1) dv_1 + (1 - F(v_{1N}^*)) \int_{p_{2N}}^1 (1 - F(v_2)) dv_2 \\
&+ F(v_{1N}^*) \int_{p_{1N}}^1 (1 - F(v_2)) dv_2 - F^N
\end{aligned}
$$

$$
\begin{aligned}
\Delta &= U^N(\mathbf{p^N}) - U_n^S(\mathbf{p^N}) = \\
&= (1 - F(v_{1N}^*)) \left( \int_{p_2}^1 (1 - F(v_2)) dv_2 - \int_{p_2 - \beta}^1 (1 - F(v_2)) dv_2 \right)
\end{aligned}
$$

$$
\begin{aligned}
U_n^S(\mathbf{p^N}) = &\int_{v_{1N}^*}^1 \big(v_1 - p_{1N} + \int_{p_{2N} - \beta}^1 (v_2 + \beta - p_{2N}) f(v_2) dv_2\big) f(v_1) dv_1 \\
&+ F(v_{1N}^*) \int_{p_{1N}}^1 (v_2 - p_{1N}) f(v_2) dv_2 - F^N \\
&= \int_{v_{1N}^*}^1 \big(v_1 - p_{1N}) f(v_1) dv_1 + (1 - F(v_{1N}^*)) \int_{p_{2N} - \beta}^1 (1 - F(v_2)) f(v_2) dv_2 \\
&+ F(v_{1N}^*) \int_{p_{1N}}^1 (1 - F(v_2)) dv_2 - F^N
\end{aligned}
$$

$$
\begin{aligned}
v_{1N}^* &= p1 + \int_{p_{1N}}^{p_{2N}} (1 - F(v_2)) dv_2 \\
v_{1S}^* &= p1 + \int_{p_{1N}}^{p_{2N} - \beta} (1 - F(v_2)) dv_2
\end{aligned}
$$

The profits of the Sophisticated are:

$$
\Pi^S = S^S(\mathbf{p^S}) - U^S(\mathbf{p^S})
$$

where

$$
\begin{aligned}
U^S(\mathbf{p^S}) &= U^S(\mathbf{p^N}) - U^N(\mathbf{p^N}) = \\
&= \int_{v_{1S}^*}^{v_{1N}^*} (v_1 - p_1) f(v_1) dv_1 - (F(v_{1N}^*) - F(v_{1S}^*)) \int_{p_1}^1 (1 - F(v_2)) dv_2 \\
&- (1 - F(v_{1N}^*)) \int_{p_2}^1 (1 - F(v_2)) dv_2 + (1 - F(v_{1S}^*)) \int_{p_2 - \beta}^1 (1 - F(v_2)) dv_2
\end{aligned}
$$

Thus, the profit function for the screening model is:

$$
\Pi = \gamma \Pi^S + (1 - \gamma) \Pi^N
$$

$$
\begin{aligned}
\frac{d\Pi}{dv_{1N}^*} &= f(v_{1N}^*)\Bigg((\gamma-1)(c-p_2)F(p_2-\beta)-(\gamma-1)\left(\int_{p_1}^{1}(v_2-c)f(v_2)dv_2\right)\\
&-2c(\gamma-1)+g\int_{p_1}^{1}(1-F(v_2))dv_2-\int_{p_2}^{1}(1-F(v_2))d\mathrm{v}2+\gamma(p_1+p_2)-p_2-v_{1N}^*\Bigg)\\
\frac{d\Pi}{dp_1} &= (1-\gamma)(p_1-c)f(p_1)F(v_{1N}^*)+\gamma\Bigg((F(p_1)-1)(F(v_{1N}^*)-F(v_{1S}^*))\Bigg)\\
&+\gamma\int_{v_{1S}^*}^{v_{1N}^*}f(v_{1N}^*)dv_1\\
\frac{d\Pi}{dp_2} &= F(p_2-\beta)\Bigg((1-\gamma)F(v_{1N}^*)+\gamma F(v_{1S}^*)-1\Bigg)\\
&+\gamma(F(v_{1N}^*)-F(v_{1S}^*))-F(p_2)(F(v_{1N}^*)-1)\\
&-(1-\gamma)(p_2-c)(F(v_{1N}^*)-1)f(p_2-\beta)\\
\frac{dv_{1N}^*}{dp_1} &= F(p_1)\\
\frac{dv_{1N}^*}{dp_2} &= 1-F(p_2)
\end{aligned}
$$

Then, the first order condition with respect to $p_1$ is:

$$\frac{d\Pi}{dp_1} = \frac{d\Pi}{dv_{1N}^*}\frac{dv_{1N}^*}{dp_1} + \frac{d\Pi}{dp_1} =$$

$$= f(v_{1N}^*)F(p_1)\Bigg((\gamma-1)(c-p_2)F(p_2-\beta) - (\gamma-1)\left(\int_{p_1}^{1}(v_2-c)f(v_2)dv_2\right)$$

$$- 2c(\gamma-1) + \gamma\int_{p_1}^{1}(1-F(v_2))dv_2 - \int_{p_2}^{1}(1-F(v_2))dv_2 + \gamma(p_1+p_2) - p_2 - v_{1N}^*\Bigg)$$

$$+ (1-\gamma)(c-p_1)f(p_1)F(v_{1N}^*)$$

$$+ \gamma\left((F(p_1)-1)(F(v_{1N}^*)-F(v_{1S}^*)) + \int_{v_{1S}^*}^{v_{1N}^*} f(v_1)dv_1\right) =$$

$$= f(v_{1N}^*)F(p_1)\Bigg((\gamma-1)(c-p_2)F(p_2-\beta) - (\gamma-1)\left(\int_{p_1}^{1}(v_2-c)f(v_2)dv_2\right)$$

$$- 2c(\gamma-1) + \gamma\int_{p_1}^{1}(1-F(v_2))dv_2 - \int_{p_2}^{1}(1-F(v_2))dv_2 + \gamma(p_1+p_2) - p_2 - v_{1N}^*\Bigg)$$

$$+ (1-\gamma)(c-p_1)f(p_1)F(v_{1N}^*)$$

$$+ \gamma\left((F(p_1)-1)(F(v_{1N}^*)-F(v_{1S}^*)) + (F(v_{1N}^*)-F(v_{1S}^*))\right)$$

after some algebra[14] and substituting $v_{1N}^*$:

$$\frac{d\Pi}{dp_1} = F(p_1)(1-\gamma)f(v_{1N}^*)\Big((p_2-c)F(p_2-\beta) + F(p_1)(c-p_1) + c - p_2\Big)$$

$$+ f(p_1)(1-\gamma)(c-p_1)F(v_{1N}^*) + F(p_1)\gamma(F(v_{1N}^*)-F(v_{1S}^*)) = 0$$

The first order condition with respect to $p_2$ is:

[14]
$$(1-\gamma)\int_{p_1}^{1}(v_2-c)f(v_2)dv_2 + \gamma\int_{p_1}^{1}(1-F(v_2))dv_2 - \int_{p_2}^{1}(1-F(v_2))dv_2$$
$$= -(1-\gamma)(1-F(p_1))c + (1-\gamma)(1-p_1F(p_1)) + \gamma(1-p_1) - (1-p_2) - \int_{p_1}^{p_2}F(v_2)dv_2$$

$$\frac{d\Pi}{dp_2} = \frac{d\Pi}{dv_{1N}^*}\frac{dv_{1N}^*}{dp_2} + \frac{d\Pi}{dp_2} =$$

$$= f(v_{1N}^*)(1 - F(p_2))\Bigg((g-1)(c-p_2)F(p_2-\beta) + (1-\gamma)\left(\int_{p_1}^{1}(v_2-c)f(v_2)dv_2\right)$$

$$- 2c(\gamma-1) + g\int_{p_1}^{1}(1-F(v_2))dv_2 - \int_{p_2}^{1}(1-F(v_2))dv_2 + \gamma(p_1+p_2) - p_2 - v_{1N}^*\Bigg)$$

$$+ (\gamma-1)(c-p_2)(F(v_{1N}^*)-1)f(p_2-\beta) + F(p_2-\beta)((1-\gamma)F(v_{1N}^*) + \gamma F(v_{1S}^*) - 1) = 0$$

then again after some algebra and substituting $v_{1N}^*$:

$$\frac{d\Pi}{dp_2} = (F(p_2)-1)(\gamma-1)f(v_{1N}^*)((p_2-c)F(p_2-\beta) + F(p_1)(c-p_1) + c - p_2)$$

$$+ (\gamma-1)(c-p_2)(F(v_{1N}^*)-1)f(p_2-\beta) + F(p_2-\beta)(\gamma F(v_{1N}^*) + \gamma F(v_{1S}^*) + F(v_{1N}^*) - 1)$$

$$+ \gamma(F(v_{1N}^*) - F(v_{1S}^*)) - F(p_2)(F(v_{1N}^*) - 1) = 0$$

Solving the system of the first order conditions:

$$p_2 = c + \frac{F(p_1)^2 f(v_{1N}^*)(F(p_2)-F(p_2-\beta))((\gamma(F(v_{1N}^*)-F(v_{1S}^*))+1-F(v_{1N}^*))}{(1-\gamma)f(p_1)f(v_{1N}^*)(1-F(p_2))\Big(1-F(p_2-\beta)F(v_{1N}^*)+f(p_2-\beta)(1-F(v_{1N}^*))(f(v_{1N}^*)F(p_1)^2+f(p_1)F(v_{1N}^*))\Big)}$$

$$+ \frac{f(p_1)F(v_{1N}^*)\Big(\gamma(1-F(p_2-\beta))(F(v_{1N}^*)-F(v_{1S}^*))+(1-F(v_{1N}^*))(F(p_2)-F(p_2-\beta))\Big)}{(1-\gamma)f(p_1)f(v_{1N}^*)(1-F(p_2))\Big(1-F(p_2-\beta)F(v_{1N}^*)+f(p_2-\beta)(1-F(v_{1N}^*))(f(v_{1N}^*)F(p_1)^2+f(p_1)F(v_{1N}^*))\Big)}$$

$$p_1 = c + \frac{F(p_1)\Big(\gamma(1-F(v_{1N}^*))f(p_2-\beta)(F(v_{1N}^*)-F(v_{1S}^*))\Big)}{(1-\gamma)\Big((1-F(v_{1N}^*))f(p_2-\beta)\Big(F(p_1)^2f(v_{1N}^*)+f(p_1)F(v_{1N}^*)\Big)+f(p_1)(1-F(p_2))f(v_{1N}^*)F(v_{1N}^*)(1-F(p_2-\beta))\Big)}$$

$$- \frac{F(p_1)\Big(f(v_{1N}^*)(F(p_2)-F(p_2-\beta))(1-F(p_2-\beta))((\gamma(F(v_{1N}^*)-F(v_{1S}^*))+1-F(v_{1N}^*))\Big)}{(1-\gamma)\Big((1-F(v_{1N}^*))f(p_2-\beta)\Big(F(p_1)^2f(v_{1N}^*)+f(p_1)F(v_{1N}^*)\Big)+f(p_1)(1-F(p_2))f(v_{1N}^*)F(v_{1N}^*)(1-F(p_2-\beta))\Big)}$$

Then $p_2 > c$ since $F(p_2) < F(p_2-\beta)$, $F(v_{1N}^*) > F(v_{1S}^*)$

Moreover, $p_1 < c$ if:

$$F(p_2) - F(p_2-\beta) < \frac{\gamma f(p_2-\beta)(1-F(v_{1N}^*))(F(v_{1N}^*)-F(v_{1S}^*))}{f(v_{1N}^*)(1-F(p_2-\beta)(1-F(v_{1N}^*)+\gamma(F(v_{1N}^*)-F(v_{1N}^*)))}$$

or

$$F(p_2) - F(p_2-\beta) < \frac{f(p_2-\beta)(1-F(v_{1N}^*))(F(v_{1N}^*)-F(v_{1S}^*))}{f(v_{1N}^*)(1-F(p_2-\beta))(1-F(v_{1S}^*))}$$

and

$$\gamma < \frac{(F(p_2) - F(p_2 - \beta))f(v_{1N}^*)(1 - F(v_{1N}^*))(1 - F(p_2 - \beta))}{(F(v_{1N}^*) - F(v_{1S}^*))((1 - F(v_{1N}^*))f(p_2 - \beta) - (F(p_2) - F(p_2 - \beta))f(v_{1N}^*)(1 - F(p_2 - \beta)))}$$

### Assuming Uniform Distribution

The maximization problem of the consumer becomes:

$$\max_{\mathbf{p^N}} \quad \Pi = \gamma\Big( - (U^S(\mathbf{p^N}) - U^N(\mathbf{p^N}))\Big) + (1-\gamma)\Big(S^N(\mathbf{p^N}) + \Delta\Big)$$

where

$$S^N(\mathbf{p^N}) = \int_{v_{1N}^*}^1 \big(v_1 - c + \int_{p_{2N}-\beta}^1 (v_2 + \beta - c)dv_2\big)dv_1 + v_{1N}^* \int_{p_{1N}}^1 (v_2 - c)dv_2$$

$$U^S(\mathbf{p^N}) = \int_{v_{1S}^*}^1 \big(v_1 - p_{1N} + \int_{p_{2N}-\beta}^1 (v_2 + \beta - p_{2N})dv_2\big)dv_1 + v_{1S}^* \int_{p_{1N}}^1 (v_2 - p_{1N})dv_2 - F^N$$

$$U^N(\mathbf{p^N})) = \int_{v_{1N}^*}^1 \big(v_1 - p_{1N} + \int_{p_{2N}}^1 (v_2 - p_{2N})dv_2\big)dv_1 + v_{1N}^* \int_{p_{1N}}^1 (v_2 - p_{1N})dv_2 - F^N$$

$$\Delta = U^N(\mathbf{p^N}) - \tilde{U}(\mathbf{p^N})$$

$$\tilde{U}(\mathbf{p^N}) = \int_{v_{1N}^*}^1 \big(v_1 - p_{1N} + \int_{p_{2N}-\beta}^1 (v_2 + \beta - p_{2N})dv_2\big)dv_1 + v_{1N}^* \int_{p_{1N}}^1 (v_2 - p_{1N})dv_2 - F^N$$

$$v_{1N}^* = p1 + \int_{p_{1N}}^{p_{2N}} (1 - v_2)dv_2$$

$$v_{1S}^* = p1 + \int_{p_{1N}}^{p_{2N}-\beta} (1 - v_2)dv_2$$

Making the calculation we take:

$$\begin{aligned}
\max_{p_1,p_2} \Pi = {} & \frac{1}{8}(8 - 8\gamma - (\beta)^4\gamma + 4(\beta)^3\gamma(p_2 - 1) + 2(\beta)^2\gamma(p_1^2 - 3p_2^2 + 2p_2 - 4) \\
& + 4\beta(\gamma - p_2)(p_1^2 - 2 - p_2^2 + 2p_2) - 4c(\gamma - 1)(4(p_2 - 1) \\
& + (p_1 - p_2)(p_1 + p_1^2 - p_2^2 - 3p_2) + \beta(-2 + p_1^2 - (-2 + p_2)p_2)) \\
& + (\gamma - 1)(3p_1^4 + 2p_1^2(4 - 3p_2)p_2 + p_2^2(8 + p_2(-8 + 3p_2))))
\end{aligned}$$

The derivative with respect to $p_1$:

$$\frac{d\Pi}{dp_1} = \frac{1}{2}\Bigg(p_1\left(\gamma\beta^2 + 2\beta(\gamma - p_2) + (\gamma-1)\left(3p_1^2 + (4-3p_2)p_2\right)\right)$$
$$- c(\gamma-1)\left(2p_1(\beta - p_2 + 1) + 3p_1^2 - (p_2-2)p_2\right)\Bigg)$$

with respect to $p_2$:

$$\frac{d\Pi}{dp_2} = \frac{1}{2}\Bigg(\gamma\beta^3 - 3\gamma\beta^2(p_2-1) + c(\gamma-1)\big(2\beta(p_2-1) + p_1^2$$
$$+ 2p_1(p_2-1) - 3(p_2-2)p_2 - 4\big) - \beta\left(2\gamma(p_2-1) + p_1^2 + (4-3p_2)p_2 - 2\right)$$
$$- (\gamma-1)\left(p_1^2(3p_2-2) + p_2(-3(p_2-2)p_2 - 4)\right)\Bigg)$$

## Cost equals to zero

Let assume that the cost is zero, $c = 0$, then the first order conditions are:

$$\frac{d\Pi}{dp_{1N}} = \frac{1}{2}p_{1N}\left(\beta^2\gamma + 2\beta(\gamma - p_{2N}) + (\gamma-1)\left(3p_{1N}^2 + (4-3p_{2N})p_{2N}\right)\right)$$
$$\frac{d\Pi}{dp_{2N}} = \frac{1}{2}\big(\beta^3\gamma - 3\beta^2\gamma(p_{2N}-1) - \beta\left(2\gamma(p_{2N}-1) + p_{1N}^2 + (4-3p_{2N})p_{2N} - 2\right)$$
$$- (\gamma-1)\left(p_{1N}^2(3p_{2N}-2) + p_{2N}(-3(p_{2N}-2)p_{2N} - 4)\right)\big)$$

then, the optimal price for the naive consumer of the first and the second unit are :

$$p_{1N} = 0$$
$$p_{2N} = \frac{1}{6(\gamma-1)}\left(2^{2/3}A + \frac{2\sqrt[3]{2}\gamma\beta((\gamma-1)(3\beta+2)+\beta)}{A} - 2\beta + 4\gamma - 4\right)$$

,where $A$ is:

$$A = \sqrt[3]{-2\beta^3 + \sqrt{\left(\beta^3\left(9(g-1)g^2+2\right) + 3\beta^2\gamma(3\gamma-1)(\gamma-1) + 6\beta(\gamma+1)(\gamma-1)^2 + 8(\gamma-1)^3\right)^2 - 4\beta^3((\gamma-1)\gamma(3\beta+2)+\beta)^3} - \gamma^3(3\beta(3\beta(\beta+1)+2)+8) + 3\gamma^2(\beta(\beta(3\beta+4)+2)+8) - 3\gamma(\beta(\beta-2)+8) - 6\beta + 8}$$

Examining the first order conditions at $p_{1N}, p_{2N} = \{0, 0\}$ we see that:

$$\left.\frac{d\Pi}{dp_{1N}}\right|_{\{0,0\}} = 0, \quad \left.\frac{d\Pi}{dp_{1N}}\right|_{\{p_{1N}=0\}} = 0$$
$$\left.\frac{d\Pi}{dp_{2N}}\right|_{\{0,0\}} = \frac{1}{2}\left(\beta(2+2\gamma) + 3(\beta)^2\gamma + (\beta)^3\gamma\right) \geq 0$$

Thus the equilibrium is $\mathbf{p_N} = \{p_{1N} = 0, p_{2N} > c, F^N = U^N(\mathbf{p_N})\}$ and $\mathbf{p_S} = \{p_{1S} = 0, p_{2S} = 0, F^S = U^S(\mathbf{p_S}) - F^S - (U^S(\mathbf{p_N}) - U^N(\mathbf{p_N}))\}$

**Fixed Fee of Sophisticated** :

The fixed fee of the Sophisticated consumer can be derived from her incentive compatibility constraint thus it is:

$$F^S = \int_{v_{1S}^*}^{1} \big(v_1 - p_{1S} + \int_{p_{2S}-\beta}^{1} (v_2 + \beta - p_{2S})dv_2\big)dv_1 + v_{1S}^* \int_{p_{1S}}^{1} (v_2 - p_{1S})dv_2$$
$$- \left(\int_{v_{1S}^*}^{1} \big(v_1 - p_{1N} + \int_{p_{2N}-\beta}^{1} (v_2 + \beta - p_{2N})dv_2\big)dv_1 + v_{1S}^* \int_{p_{1N}}^{1} (v_2 - p_{1N})dv_2 - F^N\right)$$

,where

$$F^N = \int_{v_{1N}^*}^{1} \big(v_1 - p_{1N} + \int_{p_{2N}}^{1} (v_2 - p_{2N})dv_2\big)dv_1 + v_{1S}^* \int_{p_{1N}}^{1} (v_2 - p_{1N})dv_2$$

thus since $p_{1S} = p_{2S} = c = 0$, $p_{1N} = 0$ and $p_{2N} > 0$ then the above equation becomes:

$$F^S = \frac{1}{4}\left(4 + \beta(2 + \beta + 2(4 + \beta(3+\beta))p_{2N} - 3(2+\beta)p_{2N}^2 + 2p_{2N}^3)\right)$$

checking numerically substituting the prices it is reasonable expect for the parameter levels that there is not real root.

**Incentive Compatibility Constraint of the Naive** : In order to show that the constraint that was relaxed, it really slacks at the optimum, it is needed to show that:

$$U^N(\mathbf{p^N}) > U^N(\mathbf{p^S})$$

thus since at the equilibrium $U^N(\mathbf{p^N}) = 0$ and the expected utility of the naive consumer at $\mathbf{p^S}$ equals 1 then it needs to be shown that:

$$0 > 1 - F^S \Rightarrow F^S > 1$$

which is true for $0 \leq \beta \leq 1$ and $p_2 > 0$.

# 8 Appendix B: Partially Naive Consumers

The optimization problem of the firm is:

$$\max_{U^*, p_1, p_2} \Pi = S^S - U^P + (U^P - \tilde{U}) = S - U^P - \Delta \quad \text{s.t.} \quad U^N \geq 0$$

and optimal consumption rule is:

$$v_{1P}^* = p_1 + \int_{p_1}^{p_2 - \tilde{\beta}} (1 - F(v_2))dv_2$$

The expected gross surplus is the one produced in a market with a habit forming consumer.

$$S = \int_{v_{1P}^*}^1 (v_1 - c)dF(v_1) + F(v_{1P}^*) \int_{p_1}^1 (v_2 - c)dF(v_2) + \int_{v_{1P}^*}^1 \int_{p_2 - \beta}^1 (v_2 + \beta - c)f(v_2)dv_2 dF(v_1)$$

Moreover, $\Delta$ is the difference between the true utility and the perceived of the consumer.

$$\begin{aligned}
\Delta &= \tilde{U}(\mathbf{p^N}) - U^N(\mathbf{p^N}) = \\
&= \int_{v_{1P}^*}^1 \left( v_1 - p_1 + \int_{p_2 - \beta}^1 (v_2 + \beta - p_2)dF(v_2) \right) dF(v_1) \\
&+ F(v_{1P}^*) \int_{p_1}^1 (v_2 - p_1)dF(v_2) - F^N \\
&- \left( \int_{v_{1P}^*}^1 \left( v_1 - p_1 + \int_{p_2 - \tilde{\beta}}^1 (v_2 + \tilde{\beta} - p_2)dF(v_2) \right) dF(v_1) \right. \\
&\left. + F(v_{1P}^*) \int_{p_1}^1 (v_2 - p_1)dF(v_2) - F^N \right)
\end{aligned}$$

Thus, $\Delta$ is:

$$\Delta = (1 - F(v_{1P}^*)) \left( \int_{p_2-\beta}^{p_2-\tilde{\beta}} (1 - F(v_2))dv_2 \right)$$

Then the first order conditions with respect to $p_1$ is:

$$\frac{\partial \Pi}{\partial p_1} = \frac{\partial S}{\partial v_{1P}^*} \frac{\partial v_{1P}^*}{\partial p_1} + \frac{\partial \Delta}{\partial v_{1P}^*} \frac{\partial v_{1P}^*}{\partial p_1} + \frac{\partial \Delta}{\partial p_1} + \frac{\partial S}{\partial p_1}$$

$$\begin{aligned}
\frac{\partial S}{\partial v_{1P}^*} &= \left( -v_{1P}^* + c - \int_{p_2-\beta}^{1} (v_2 + \beta - c)dF(v_2)) + \int_{p_1}^{1} (v_2 - c)dF(v_2)) \right) f(v_{1P}^*) \\
\frac{\partial \Delta}{\partial v_{1P}^*} &= - \left( \int_{p_2-\beta}^{p_2-\tilde{\beta}} (1 - F(v_2))dv_2 \right) f(v_{1P}^*) \\
\frac{\partial v_{1P}^*}{\partial p_1} &= 1 - (1 - F(p_1)) = F(p_1) \\
\frac{\partial S}{\partial p_1} &= -F(v_{1P}^*)(p_1 - c)f(p_1)
\end{aligned}$$

Then the first order condition is:

$$\begin{aligned}
\frac{\partial \Pi}{\partial p_1} &= \left( -v_{1P}^* + c - \int_{p_2-\beta}^{1} (v_2 + \beta - c)dF(v_2)) + \int_{p_1}^{1} (v_2 - c)dF(v_2)) \right) f(v_{1P}^*)F(p_1) \\
&+ \left( \int_{p_2-\beta}^{p_2-\tilde{\beta}} (1 - F(v_2))dv_2 \right) f(v_{1P}^*)F(p_1) - F(v_{1P}^*)(p_1 - c)f(p_1) = \\
&= \Big( - p_1 + (p_1 - p_2 + \tilde{\beta}) + c + c(1 - F(p_2 - \beta)) - c(1 - F(p_1)) \Big) f(v_{1P}^*)F(p_1) \\
&+ \Big( - (1 - (p_2 - \beta)F(p_2 - \beta)) - \beta(1 - F(p_2 - \beta)) \Big) f(v_{1P}^*)F(p_1) \\
&+ \Big( 1 - p_1F(p_1) + (p_2 - \tilde{\beta} - p_2 + \beta) \Big) f(v_{1P}^*)F(p_1) \\
&- F(v_{1P}^*)(p1 - c)f(p_1) = 0
\end{aligned}$$

Then

$$p_1 = c - (p_2 - c) \frac{f(v_{1P}^*)F(p_1)(1 - F(p_2 - \beta))}{(f(v_{1P}^*)F(p_1)^2 + f(p_1)F(v_{1P}^*))}$$

The first order condition with respect to $p_2$ is:

$$\frac{\partial \Pi}{\partial p_2} = \frac{\partial S}{\partial v_{1P}^*} \frac{\partial v_{1P}^*}{\partial p_2} + \frac{\partial \Delta}{\partial v_{1P}^*} \frac{\partial v_{1P}^*}{\partial p_2} + \frac{\partial \Delta}{\partial p_2} + \frac{\partial S}{\partial p_2}$$

then the respective derivatives are:

$$\frac{\partial S}{\partial p_2} = (1 - F(v_{1P}^*))(-1)(p_2 - c)f(p_2 - \beta)$$
$$\frac{\partial \Delta}{\partial p_2} = 1 - F(p_2 - \tilde{\beta}) - (1 - F(p_2 - \beta)) = F(p_2 - \beta) - F(p_2 - \tilde{\beta})$$
$$\frac{\partial v_{1P}^*}{\partial p_2} = 1 - F(p_2 - \tilde{\beta})$$

Thus the first order condition with respect to $p_2$ becomes:

$$\frac{\partial \Pi}{\partial p_2} = \Bigg( - v_{1P}^* + c - \int_{p_2-\beta}^{1} (v_2 + \beta - c)dF(v_2)) + \int_{p_1}^{1} (v_2 - c)dF(v_2))$$
$$+ \int_{p_2-\beta}^{p_2-\tilde{\beta}} (1 - F(v_2))dv_2 \Bigg) f(v_{1P}^*)(1 - F(p_2 - \tilde{\beta})) + F(p_2 - \beta) - F(p_2 - \tilde{\beta})$$
$$- (1 - F(v_{1P}^*))(p_2 - c)f(p_2 - \beta) =$$
$$= (-p_1 + (p_1 - p_2 + \tilde{\beta}) + c + c(1 - F(p_2 - \beta)) - c(1 - F(p_1))$$
$$- (1 - (p_2 - \beta)F(p_2 - \beta)) - \beta(1 - (p_2 - \beta)) + 1 - p_1 F(p_1) +$$
$$+ (p_2 - \tilde{\beta} - p_2 + \beta))(v_{1P}^*)(1 - F(p_2 - \tilde{\beta})) - (1 - F(v_{1P}^*))(p_2 - c)f(p_2 - \beta)$$
$$- (F(p_2 - \beta) - F(p_2 - \tilde{\beta})) = 0$$

[15]

$$p_2 = c + (c - p_1)\frac{f(v_{1P}^*)F(p_1)(1 - F(p_2 - \tilde{\beta})}{(f(v_{1P}^*)(1 - F(p_2 - \beta))(1 - F(p_2 - \tilde{\beta})) + f(p_2 - \beta)(-1 + F(v_{1P}^*)))}$$
$$+ \frac{F(p_2 - \tilde{\beta})) - (F(p_2 - \beta)}{f(v_{1P}^*)(1 - F(p_2 - \beta))(1 - F(p_2 - \tilde{\beta})) + f(p_2 - \beta)(1 - F(v_{1P}^*))}$$

Finally solving the above system of equations we get:

$$p_1 = c - \frac{f(v_{1P}^*)F(p_1)(1 - F(p_2 - \beta))(F(p_2 - \tilde{\beta}) - F(p_2 - \beta))}{f(p_1)f(v_{1P}^*)(1 - F(p_2 - \beta))(1 - F(p_2 - \tilde{\beta}))F(v_{1P}^*) + f(p_2 - \beta)(1 - F(v_{1P}^*))(f(v_{1P}^*)F(p_1)^2 + f(p_1)F(v_{1P}^*)))}$$

[15]

$$- \int_{p_1}^{p_2} (1 - F(v_2))dv_2 - \int_{p_2}^{1} (1 - F(v_2))dv_2 - \int_{p_1}^{1} F(v_2)dv_2 = -1 + p_1$$
$$\int_{p_1}^{1} (v_2 - c)dF(v_2) = 1 - p_1 F(p_1) - \int_{p_1}^{1} F(v_2)dv_2$$

$$p_2 = c + \frac{(F(p_2 - \tilde{\beta}) - F(p_2 - \beta))(f(v_{1P}^*)F(p_1)^2 + f(p_1)F(v_{1P}^*))}{f(p_1)f(v_{1P}^*)(1 - F(p_2 - \beta))(1 - F(p_2 - \tilde{\beta}))F(v_{1P}^*) + f(p_2 - \beta)(1 - F(v_{1P}^*))(f(v_{1P}^*)F(p_1)^2 + f(p_1)F(v_{1P}^*))}$$