Discussion Paper Series – CRC TR 224

Discussion Paper No. 098
Project B 05

Optimal Non-Linear Pricing Scheme when Consumers are Habit Forming

Eleftheria Triviza*

June 2019

*Department of Economics. University of Mannheim and MaCCI. Email: etriviza@mail.unimannheim.de

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Optimal Non-Linear Pricing Scheme when Consumers are Habit Forming

Eleftheria Triviza*

June 4, 2019

Abstract

This article analyses how consumers' habit formation affects firms' pricing policies. We consider both sophisticated consumers, who realize that their current consumption will affect future consumption, and naive consumers, who do not. The optimal contract for sophisticated consumers is a two-part tariff. The main result is that under naive habit formation, the optimal pricing pattern is a three-part tariff; namely a fixed fee, with some units priced below cost — and after their end — pricing above marginal cost. This holds both under symmetric and asymmetric information. **JEL:** L11, D11, D42, D82

Keywords: three-part tariff, nonlinear pricing, naivete, habit formation

*Department of Economics. University of Mannheim and MaCCI. Email: etriviza@mail.uni-mannheim.de

I thank Vincenzo Denicoló and Francesca Barigozzi, my Phd supervisors, for all the helpful discussions and support. I thank Mark Armstrong, Sarah Auster, Giacomo Calzolari, Martin Peitz, Kathryn Spier, Emanuelle Tarantino, Tomaso Valletti and Piercarlo Zanchettin. I also thank for helpful comments and suggestions seminar audiences at University of Bologna in 2016, Mannheim University in 2017, the 2018 Workshop on Behavioral Industrial Organization, the 2017 Annual Conference of the European Association for Research in Industrial Economics (EARIE), Lancaster University Management School in 2016, the 2015 Conference on Research on Economic Theory and Econometrics (CRETE). Funding by the German Research Foundation (DFG) through CRC TR 224 is gratefully acknowledged.

1 Introduction

Three-part tariffs are prevalent in telecommunications and Information Technology (IT) markets features (Grubb, 2009, 2014). These contracts include a fixed fee, an allowance of free units, and a positive price for additional units beyond the allowance. This pricing pattern is not consistent with rational consumer behavior. This article shows that the presence of consumers with naive habit forming behavior is sufficient for the optimality of three-part tariffs. More specifically, it is sufficient that the consumers are not entirely aware of how much their past consumption would affect their current valuation for the good. Importantly, this type of behavior has been well documented in telecommunications and IT markets (Bianchi and Phillips, 2005; Park, 2005).

We solve a dynamic pricing model in which a firm sets the price at the contractual stage, and the consumers decide whether or not to buy, based on their expectation of the value of their future consumption. The consumers have two consumption opportunities within the contract period: namely, they can buy the good once or twice during the contract period, depending on their needs and their valuation of the good. At the end of the period, they make the payment.

In the benchmark model, we consider sophisticated consumers who are aware that today's consumption affects future consumption. In this case, it is optimal for the monopolist to charge a two-part tariff. Since consumers know the value of their future consumption exactly, the firm finds it optimal to maximize the consumer surplus by setting marginal prices equal to the marginal cost, and to charge a fixed fee that extracts all of the consumer surplus.

This changes when we consider naive habit-forming consumers. These consumers are unaware of their habit forming behavior at the contractual stage. The monopolist, however, can recognize that they are habit-forming. In this second case, the optimal contract offered by the firm is a three-part tariff. The firm charges a marginal price above marginal cost for high volumes, a marginal price below marginal cost for low volumes, and a fixed fee.

As an intuition, naive habit forming consumers underestimate the probability of having high demand at the contractual stage. They do not expect that they will acquire a habit and thus fail to realize that the probability of consuming the next unit of the good or service will be larger. Given this bias, the firm finds it optimal to distort marginal cost pricing by charging a marginal price above marginal cost for this and, thus, for high volumes. This resulting pattern is similar to the pricing of hyperbolic discounted leisure goods, where the underestimation of demand also arises (DellaVigna and Malmendier, 2004; Eliaz and Spiegler, 2006).

Naive habit forming consumers do not make mistakes about the probability of having low demand at the contractual stage. It is only after consuming that they experience an unexpected change in their demand (Pollak, 1975). Moreover, the naive habit forming consumers evaluate their consumption decisions sequentially, and they are forward-looking. This means that they can foresee that there will be a price change in the future and they internalize this information into their decision as to whether or not to consume in the current period. They can also foresee that they may forego utility if they consume today, and expect that the next unit will be charged differently and will possibly be more costly. Thus, the monopolist finds it optimal to charge a price below marginal cost for low volumes, since the consumers are forward-looking, with the second unit priced above marginal cost, for the reasons explained before. In this way, the probability of consuming the first unit increases and the cost of forgone future utility decreases. The firm finds it optimal to increase the probability of consuming the first unit, not only because it will lead to more future consumption but also because the firm can fully extract the surplus produced from the first unit. The consumer makes no mistake at the contractual stage for the first unit, the perceived expected utility is equal to the actual expected utility, and thus the fixed fee can fully extract the surplus.

Finally, the third part of the tariff is the fixed fee. The fixed fee is equal to the expected surplus of the consumers at the contractual stage. However, the consumers undervalue the contract offered by the monopolist at the contractual stage, because they cannot foresee that they will value the good more highly the more they consume. They participate in the market, considering themselves as non-habit forming. For this reason the firm cannot extract all of the consumer surplus actually produced with the fixed fee. The monopolist mitigates the contract undervaluation, and thus extracts as much consumer surplus possible by distorting marginal prices. The direction of the distortion of the marginal prices is as discussed. The consumers, in turn, are left with a positive *misperception rent*, as given by the difference between their true expected surplus and the surplus they mistakenly perceive at the contractual period. Consequently, the naive consumers cannot be exploited through the pricing scheme.[1]

The literature to date has focused on consumers' overconfidence as the main explanation for the use of three-part tariffs (Grubb, 2009). Overconfidence means that consumers overestimate their demand when it is low, and underestimate it when it is high. The main difference between overconfidence and habit formation is that for three part tariffs to be optimal, in Grubb (2009) both mistakes are necessary. This article shows that is sufficient that the consumers underestimate high demand.

The optimality of three-part tariffs in the presence of naive habit forming consumers is confirmed when we add competition into the picture. In Section 5, we consider a market where two firms compete on a Hotelling line. The optimal pricing scheme is the same as in the monopolistic case. The only part of the tariff that differs is the fixed fee, which decreases as the market becomes more competitive.

In Section 6, we relax the assumption of full information and study the pricing strategy of a monopolist when the firm cannot observe the consumer type. We study

[1] We use the notion of exploitation according to Eliaz and Spiegler (2006), where "An exploitative contract extracts more than the agent's willingness to pay, from his first-period perspective".

the optimal screening of habit-forming consumers with differing degrees of sophistication. The monopolist knows that all consumers in the market are habit forming with different levels of sophistication and thus with different expected demands. However, the monopolist cannot observe the type of the consumer, namely whether she is sophisticated or naive.

We contend that frequently observed contract menus comprise both two and three-part tariffs and can be explained by the presence of consumers with differing levels of sophistication. We show that the firm offers both a two-part and a three-part tariff to screen between sophisticated and naive consumers. Moreover, we show that a three-part tariff is still the optimal contract for naive consumers, even when there is asymmetric information.

To understand why, consider that the sophisticated consumers would have an incentive to mimic the naive consumers. Even if they know that they are more likely to consume in the future, they would choose a contract that penalizes large consumption levels with high marginal prices. The incentive to mimic arises because the contract made for naive consumers charges a fixed fee that does not extract all of the consumers surplus. Thus, the sophisticated consumers — by mimicking the naive consumer — would be left with a positive rent ex-post. For this reason, the optimal contract for sophisticated consumers charges the same marginal prices as in the full information case, but a smaller fixed fee. The firm does not extract all of the consumer surplus to make the contract incentive compatible, and the sophisticated consumers are left with an information rent. Consequently, the presence of naive consumers in the market exerts a positive externality on the sophisticated consumers.

The optimal contract for naive consumers is still a three-part tariff as in the full information case because of the same economic forces. The difference now is that the contracts should be incentive compatible and not attractive for the sophisticated consumer. For this reason, we observe an increase in the marginal prices per unit.

The firm still cannot exploit the naive consumers who are left with a positive but smaller, in this case, misperception rent. This is because the increase in the marginal prices decreases the probability of having high demand and therefore it decreases the consumer surplus that could be produced and not fully extracted. Thus, the naive consumers are worse off when there are sophisticated consumers in the market.

Even if the optimal marginal price for low volumes is smaller than the marginal cost, the naive consumer underconsumes compared to the sophisticated one. Underconsumption happens because the consumer is forward-looking. She considers the price change and the potentially lost second consumption opportunity when deciding whether or not to consume the first unit without considering its benefit. The firm, on the other hand, finds it optimal to charge such a pricing scheme, making second unit consumption less likely. The consumer cannot anticipate its real value and the firm cannot fully extract its surplus. Thus, the firm needs to mitigate the undervaluation of the contract and the profit losses, charging a contract that makes it less likely to be in this situation. Similarly, for high volumes, the consumer also underconsumes. The firm in both periods exacerbates the mistake the consumer makes with respect to being sophisticated. This underconsumption leads to deadweight loss and thus to an inefficient allocation. A policy that informs the consumers about their habit forming behavior would help to avoid this inefficiency.

The article proceeds as follows. Section 2 discusses the related literature. Section 3 is devoted to the model setup, and Section 4 presents the case of full information with a monopolist in the market. Section 5 considers the case of an oligopolistic market. Section 6 discusses the case of asymmetric information. Finally, Section 7 summarizes and concludes.

2 Literature

This article is related to different streams of the literature. First, it is related to models that 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 costs would explain the use of a three-part tariff. He claims that a three-part tariff is the optimal pricing scheme when the behavior of the consumer is characterized by an overestimation of the demand, when the demand is low, and an underestimation of the demand when it is high. In our case, we propose a different behavior that could explain this pricing scheme without both mistakes necessarily being present. Moreover, we study an environment in which the firm observes the amount actually consumed by the consumer in each period,[2] and not only the amount the consumer has bought.

Grubb (2014) shows that inattentive behavior has similar features and implications to overconfident behavior. The common element between our model and Grubb (2014) model is that we both consider the consumption dynamically within the contract period; however, we propose a different type of behavior.

Eliaz and Spiegler (2008) consider a model where consumers have biased priors. We also do this, but with only two types of ex-post demand: high or low. The consumers are optimistic and think that the good state is more likely to happen. They describe a situation where consumers are dynamically inconsistent, and they under or overestimate average demand. Thus, Eliaz and Spiegler (2008) study an entirely different behavioral bias, having only the biased priors in common.

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 goods (Becker et al., 1991;

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

Driskill and McCafferty, 2001) are connected to our study. However, we consider habit formation and optimal pricing within a contract period, where the firm cannot renegotiate the price during the contract period.

Moreover, the discussion of a naive habit forming consumer is closely related to the articles that consider the optimal nonlinear pricing induced by various types of consumers' biases or nonstandard preferences.[3]

On the one hand, there are articles discussing biased beliefs, such as naive quasi-hyperbolic discounting for leisure goods (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 consequence of these behavioral biases is an underestimation of the demand, which results in marginal prices above marginal cost. 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 an overestimation of demand may explain prices below marginal costs, but not above.

DellaVigna and Malmendier (2004) were the first to point out that firms may design contracts to exacerbate consumer's mistakes. Since their pioneering contribution, many articles have explored the specific way of exploiting consumer naivety. In our model, the firm offers a contract that exacerbates a consumer's mistake but cannot extract all of 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. There are two kinds of consumers' mistakes more often analyzed in the literature. Firstly, the consumer

[3]For standard, rational preferences, see Mussa and Rosen (1978) and Maskin and Riley (1984). They explain contracts with high marginal prices for early units and marginal cost pricing for late units consumed; although, they cannot predict the inverse which is the main characteristic of three part tariffs.

does not understand all of the features of a contract (all prices and fees) (Gabaix and Laibson, 2006; Armstrong and Vickers, 2012). Another kind of mistake is to mispredict their own behavior concerning the product (DellaVigna and Malmendier, 2004). The latter kind of mistake is closer to the model we study here, and 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). In contrast to this literature, the optimality of the pricing scheme is not the result of a screening mechanism.

This section is also related to the literature on sequential screening of consumer with standard preferences. In these models the consumers know at the contracting period the distribution of their valuation for the good and subsequently they learn their realized valuation (Courty and Hao, 2000; Miravete, 2005; Inderst and Peitz, 2012).

Eliaz and Spiegler (2006) show that the ultimate source of gains for the principal is the non-common prior assumption. The consumer is uncertain as to whether or not her preference will change, but she knows exactly what they could change into. The firm, though, knows that the consumer's preference will change, and takes advantage of its superior information by also contracting the event that the consumer thinks unlikely to happen. The difference in their prior expectations 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. This feature becomes important because in both cases the contract is signed before the consumer experiences the change in her utility and cannot be renegotiated afterward.

3 Model Setup

This section presents the basic structure of the model. We consider a model that follows Grubb (2014) in modeling a consumer who has two consumption opportunities: one per period, and in each period purchases at most one unit of the good. Moreover, the consumers are habit-forming with differing levels of 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^θ (the price of the first unit consumed), p_2^θ (the price of the second unit consumed), and F^θ (a fixed payment). The first unit has the same price, irrespective of the period t when consumed. Time-dependent pricing would require that the firm could observe and record the opportunities to consume, as if, for example, the consumer had 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 decides to consume or not.[4] At each consecutive period $t \in \{1, 2\}$, the consumer learns the realization of a taste shock v_t, randomly drawn from a cumulative distribution function $F(v)$ with support [0,1], the same for all types of consumers and for both periods. Then, v_t is the valuation that a unit of good has in period t. Then, given her valuation, she has 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. The result of increasing marginal pricing 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 in Figure 1.

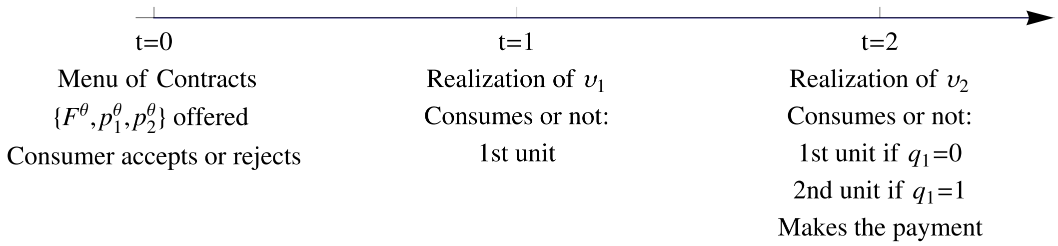

Figure 1: Timing of the game

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

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

where $\boldsymbol{\upsilon} \in [0, 1]^2$.

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

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

The expected profits per consumer are equal to 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:

$$\Pi = E\left[\left(p^\theta(\mathbf{q}(\boldsymbol{\upsilon}; \mathbf{p}), \mathbf{p}^\theta) - c\right) \mathbf{q}(\boldsymbol{\upsilon}; \mathbf{p}^\theta)\right].$$

Finally, the expected social surplus is:

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

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

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

Therefore, 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 habit forming the consumer is, and 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 the θ, the less naive is the consumer; thus, the more she realizes that she is affected by her previous consumption. Thus, $\theta = 1$ means that she is a sophisticated habit forming consumer, $0 < \theta < 1$ means that she is partially naive, and $\theta = 0$ that she is completely naive.

Every time that the consumer faces a consumption decision, she compares the valuation of the unit with her reservation price. The reservation price in each period is the threshold above which the valuation should be for the consumer to be optimal 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. This consumption rule is different for each potential consumption decision, since the decisions are taken sequentially. Thus, these thresholds are the consumption strategy of the consumer, namely the argument of maximization of her expected utility for the respective unit and period.

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 it is optimal for them to consume. During 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:

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

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

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, the optimal threshold decreases. Moreover, the first-period threshold increases if the second unit marginal price increases and decreases the more habit forming the consumer is.

Naive Habit Forming Consumer ($\theta = 0$): In period 2, the actual optimal threshold is the same as for a sophisticated consumer. 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, because 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:

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

The 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 even if the two types of consumers —

sophisticated and naive — are ex-post identical, they differ in the ex-ante perception of how habit forming they are.

In fact, the true ex-ante utility of the consumer differs from what she expects at the contractual stage. Mainly, this holds in the second period. In the first period, the optimal threshold and the expected utility are the ones of a non-habit forming consumer. Thus, the actual ex-ante expected utility for the naive consumer is:

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

Therefore, she believes that her expected utility is U^N (equation (2)), even though her actual expected utility, and the one that the firm expects that she will have, is $\tilde{U}$ (equation (3)). This whole analysis also holds when the consumer is partially naive. Namely, the consumer knows that she is habit forming, but she believes that she is less habit forming than she is.[6]

The consumer uses in period one the same threshold as she expected to use when she chose her contract in the contractual stage. More specifically, the probability of consuming in the first period is $1 - F(v^*_{1N})$, as the consumer expected at period 0. Thus, there is no difference between the consumer's expectations about her future self and how she actually acts. This means that there is no mistake that the firm could take advantage of. However, her naivety has no direct implication on the first unit marginal price.[7] The only implication of the consumer's naivety, in the first period, is related to the expected consumer surplus, which is smaller than the one that would be produced if the consumer were sophisticated.

[6] 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 is high. See the Appendix.

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

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, and 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. Hence, 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. *Let π be the actual probability of consumption and $\tilde{\pi}$ the perceived probability of consumption at the contacting period. A naive habit forming consumer makes no mistake if her demand is low, namely $\pi(q_1 = 1) = \tilde{\pi}(q_1 = 1)$ and $\pi(q_2 = 1|q_1 = 0) = \tilde{\pi}(q_2 = 1|q_1 = 0)$. Moreover, she underestimates her future demand when it is high, $\pi(q_2 = 1|q_1 = 1) > \tilde{\pi}(q_2 = 1|q_1 = 1)$.*

In addition, because of her naivety, the consumer under-evaluates the overall value of 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 β additional valuation, and she does not consider it in the ex-ante valuation of 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.

4 Informed Monopolist

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 the production of one unit of the good is $c \in (0, 1)$.

The maximization problem of the firm is:

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

This 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 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 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 extracts all the consumer surplus (see Appendix A).

Naive Consumer

Let us 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 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 is aware 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 the firm 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:

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

where $\Delta(\mathbf{p^N})$ 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)). Moreover, the firm chooses a pricing scheme that makes the participation constraint bidding, $U^N(\mathbf{p^N}) = 0$. As mentioned before, the consumer undervalues the value of the whole contract at the contracting stage, and she is not willing to pay as much as it is the actual consumer surplus produced. Thus, there is a positive rent $\Delta(\mathbf{p^N})$ that is left to the consumer. It follows that the firm cannot extract all the consumer surplus. After some simplifications, $\Delta(\mathbf{p^N})$ can be rewritten as:

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

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

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

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

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

$$
\begin{aligned}
c = 0: &\quad p_1^N = 0, \quad p_2^N > c \\
c > 0: &\quad p_1^N < c, \quad p_2^N > c,
\end{aligned}
$$

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

Proof: See Appendix A

The optimal pricing scheme when the consumer is naive habit forming resembles the scheme we observe in several markets, namely a *three part tariff*. This consist of a fixed fee, an included allowance of units, for which the marginal price equals zero, and a positive marginal price for units beyond the allowance.[8] 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 higher than the marginal cost.

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 misperceives her expected utility, the participation constraint is biased. The firm cannot extract the surplus produced through a fixed fee, since the perceived surplus is smaller than the one produced. Therefore, the firm needs to distort the marginal prices by choosing the ones that maximize $S^N - \Delta$ rather than S^N.

The exact way in which the marginal prices are distorted depends on the characteristics of the consumer's behavior. Firstly, the consumer underestimates the probability of consuming the second unit. This underestimation makes it optimal for the firm to charge a price greater than the marginal cost.[9] 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 to constrict the downward bias in consumption. However, more importantly, these marginal prices exacerbate the mistake[10] of the second unit consumption.

[8]We could consider $\beta < 0$. Think of "novelty thrill" or a "fashion good", the less novel or fashionable feels the less someone consumes it. Then, the purchasing probability is decreasing without being aware of it ex ante. In this case, it would have the opposite pricing scheme, i.e. $p_1^N < c$, $p_2^N > c$.

[9]Similar result to the behaviors of hyperbolic discounting and myopia.

[10]DellaVigna and Malmendier (2004) were the first to point out that firms might design contracts to exacerbate consumer's mistakes. Since their pioneering contribution, many articles have explored the specific ways to exploit consumer naivety.

A distorted price below marginal cost makes the consumption of the first unit more probable than marginal cost pricing. Then, it increases the probability of having a second consumption opportunity and consequently the probability of consuming two units of the good. Thus, the consumption of the second unit becomes more probable, not only because the consumer acquires a habit that she does not expect, but also because the first unit marginal price facilitates it.

Even if it seems that for the first unit there would be over-consumption, the fact that the consumer is forward-looking and naive of her habit forming behavior produces the opposite 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 under-consumption compared to the efficient allocation of the sophisticated habit forming consumer. Moreover, the larger the habit formation coefficient β, the greater the first unit threshold v_{1N}^*. The reason is that the difference between the first and second unit optimal marginal price is larger, the larger the β.

The reason for under-consumption in the first period is that in this way the firm mitigates the undervaluation of the contract and thus decreases the Δ, namely the part of the consumer surplus that cannot extract with the fixed fee.

In the profit function of the firm, there are two opposite effects. On the one hand, the firm wants to decrease the price below marginal cost that would make the consumption of the first unit more probable, and consequently, of the second unit. The mistaken expectations about the second unit are the ones that could be taken advantage of. On the other hand, the firm has an incentive to increase the first-period optimal threshold to minimize Δ, that enters negatively into its profit function. Even if the firm can overcharge the second unit because of the mistaken beliefs, it cannot extract ex-ante all of 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 under-consumption of the second unit. The optimal second unit threshold for the naive consumer is always greater than that 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 it 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 misperception rent would give an incentive to the consumer to remain naive and not pay the cost of becoming sophisticated and learning her true type. Remaining naive is beneficial for her both because the firm cannot extract all of her surplus and because she avoids paying any information cost to become sophisticated.

A typical critic is that naivety goes away with learning or can be mitigated when appropriate feedback is provided (Bolger and Önkal-Atay, 2004), although, the consumers may learn slowly (Grubb and Osborne, 2015), or forget what is learned (Agarwal et al., 2013). Moreover, three-part tariffs are optimal even when consumers are partially naive (see Appendix B), which could resemble the period in which she learns her true type.

5 Informed Duopolists

In this section, we introduce competition to examine whether or not the pricing structure that is optimal under a monopoly would also be 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 endpoints 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.[11] The competition is in the utility space and τ is the transportation cost. Moreover, S_i^N is the social surplus, U_i^N is the mistakenly expected consumers' surplus, and Δ_i the difference between the actual and the mistakenly expected consumers' surplus that is 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 τ, 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 assume that the above inequality holds and thus there is full market coverage in this market, then we have Proposition 2.

Proposition 2. *Hotelling Duopoly: Let τ be sufficiently small for strict full market coverage and the consumer be naive habit forming. Then, the optimal pricing scheme is:*

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

The marginal prices equal the monopolistic one. The competition among the firms is in the utility space, and thus it affects only the fixed fee charged as compared to

[11] 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 contract offered by the monopolist. The pricing scheme is the same in both cases because there is full market coverage. 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 τ, the less scope there is for price discrimination.

6 Uninformed Monopolist

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 γ.

The screening is done with respect to the pricing scheme.[12] The firm offers a menu of contracts. Without any loss of generality, we can restrict the analysis to the case in which it offers as many contracts as the number of types; thus, two. 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.

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

[12] We use the taxation principle because it is 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.

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*: that is, each type should not have any incentive to mimic the other at the optimal allocation. 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, the profit of the firm is greater when there is only a sophisticated consumer in the market.[13] 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 of the consumer 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 of the consumer surplus at the contracting stage.

The above discussion suggests that we expect the incentive compatibility constraint of naive, IC_N, not to bind at 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. 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 compatibility 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:

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

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:*

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

Proof: See Appendix A

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

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

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 naivet;, that is, 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* Δ 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 screening contract that the firm offers to 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 the fraction of sophisticated consumers γ is relatively small, and the fixed fee, F^N, equals the perceived consumer surplus of the naive consumer.*

Proof: See Appendix A

The firm offers a menu of contracts 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 quite 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 slackens at the optimum. This is shown in the Appendix A.

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.[14]

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

[14]See Appendix A for detailed calculations.

consumers. However, it can be shown that overall, this type of consumer is worse off in the presence of the sophisticated ones. 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 than in the full information case, and thus the naive consumer is worse off.

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.

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: $v^*_{1N}(\mathbf{p^N}) > v^*_{1N}(\mathbf{p^S}) = v^*_{1N}(\mathbf{c})$. On the other hand, the sophisticated consumer is more likely to consume when she chooses the contract made for the naive consumer, $v^*_{1N}(\mathbf{p^N}) > v^*_{1S}(\mathbf{p^N})$, and even more likely when she chooses the contract tailored for her, $v^*_{1S}(\mathbf{p^N}) > v^*_{1S}(\mathbf{p^S})$. Moreover, the more habit forming the consumer, the greater the exacerbation of the mistake, thus the more under-consumption in the first period.

Importantly, under-consumption leads to deadweight loss. There is allocative inefficiency and thus a loss of economic efficiency. The consumer is left with a positive consumer surplus and this could be seen as a reason for no policy intervention. Though, the deadweight loss that under-consumption creates, could raise concerns. Regulatory authorities may consider the need for analysis of the possible policies that could alleviate this efficiency loss. For example, a possible intervention could be to inform consumers of their habit forming behaviors.

7 Conclusion

During the last decades, the provision of a menu of contracts consisting of two-part and three-part tariffs has become prevalent in a number of markets. Moreover, there is evidence that the consumption of communication services, such as cell phones and internet, is habit forming.

This article shows that habit forming behavior can explain these observed pricing schemes. In particular, naive habit formation by consumers makes it optimal for the firm to charge a "three-part tariff".

We show that this pricing scheme is optimal if three conditions are met: (i) the consumption choice is made sequentially within the contract period; (ii) the consumer undervalues the offered contract at the contracting period; and (iii) the consumer underestimates high demand.

This explanation can be viewed as an alternative channel to the overconfidence model of Grubb (2009) that also explains this type of pricing scheme. We show that if the elements mentioned before are present, it is sufficient the consumer underestimates high demand for the introduction of the three-part tariff to be optimal; in contrast to Grubb (2009) she does not need to overestimate low demand.

Interestingly, the firm cannot take advantage of the naivety of its client. We may expect that a consumer that forms a habit without knowing could be exploited. To the contrary, the firm is worse off when it encounters a naive habit forming consumer as opposed to a sophisticated one. This is because the firm cannot charge a fixed fee that extracts all of 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 worse off.

Moreover, this article claims that the observed menu of contracts could be explained by the existence of habit forming consumers with varying levels of sophistication about their habit forming behavior. It is shown that the firm finds it optimal to offer a two-part tariff to sophisticated consumers and a 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. The sophisticated consumer has an incentive to pretend to be the naive consumer and choose the contract intended for them. It seems counter-intuitive because this contract charges highly for the high consumption, and the sophisticated consumer knows that she will form a habit that will make high consumption more likely. Though, we see that she finds it optimal to mimic the naive consumer and consume less, because in this way she is left with a rent. For this reason, the firm finds it optimal to leave information rent to the sophisticated consumers.

Hence, both types are left with a rent, since, in addition, the naive consumers are left with a mis-perception rent. Nevertheless, 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 in the marginal prices, and thus a decrease in the ex post misperception rent.

The presence of naive habit forming consumers in the market and three-part tariffs induce allocative inefficiency. The naive consumer under-consumes both units; namely the probability of consuming in both periods is less than if she was sophisticated.

Thus, there are serious welfare implications, and the need for a policy intervention to decrease the deadweight loss created seems requisite. A potential policy that could increase the overall welfare in the market would be to inform naive consumers of their habit forming behavior.

8 References

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

Armstrong, M. and J. Vickers. "Competitive price discrimination". *RAND Journal of Economics*, Vol. 32 (2001), pp. 579-605.

Armstrong, M. and J. Vickers. "Consumer protection and contingent charges". *Journal of Economic Literature*, Vol. 50 (2012), pp. 477-493.

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

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

Bolger, F. and D. Onkal-Atay. "The effects of feedback on judgmental interval predictions". *International Journal of Forecasting*, Vol. 20 (2004), pp. 29-39.

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

Courty, P. and Hao, L." Sequential screening". *The Review of Economic Studies*, Vol. 67 (2000), pp.697-717.

DellaVigna, S. and U. Malmendier. "Contract design and self-control: Theory and evidence". *The Quarterly Journal of Economics*, Vol. 119 (2004), pp. 353-402.

Driskill, R. and S. McCafferty, "Monopoly and oligopoly provision of addictive goods". *International Economic Review*, Vol. 42 (2001), pp. 43-72.

Eliaz, K. and R. Spiegler. “Contracting with diversely naive agents”. *The Review of Economic Studies*, Vol. 73 (2006), pp. 689-714.

Eliaz, K. and R. Spiegler. “Consumer optimism and price discrimination”. *Theoretical Economics*, Vol. 3 (2008), pp. 459-497.

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

Fethke, G. and R. Jagannathan. “Habit persistence, heterogeneous tastes, and imperfect competition”. *Journal of Economic Dynamics and Control*, Vol. 20 (1996), pp. 1193-1207.

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

Grubb, M. “Selling to overconfident consumers”. *The American Economic Review*, Vol. 99 (2009), pp. 1770-1807.

Grubb, M. D. “Consumer inattention and bill-shock regulation”. *The Review of Economic Studies*, Vol. 82 (2014), pp. 219-257.

Grubb, M. D. and M. Osborne. “Cellular service demand: Biased beliefs, learning, and bill shock”. *American Economic Review*, Vol. 105 (2015), pp. 234-271.

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

Heidhues, P. and B. Koszegi. “Exploiting naivete about self-control in the credit market”. *The American Economic Review*, Vol. 100 (2010), pp. 2279-2303.

Herweg, F. and K. Mierendorff . “Uncertain demand, consumer loss aversion, and at rate tariffs". *Journal of the European Economic Association*, Vol. 11 (2013), pp. 399-432.

Inderst, R. and M. Peitz. "Informing Consumers about their own Preferences". *International Journal of Industrial Organization*, Vol. 30 (2012), pp. 417-428.

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

Miao, C.-H." Consumer myopia, standardization and aftermarket monopolization". *European Economic Review*, Vol. 54 (2010), pp. 931-946.

Miravete, Eugenio J. "The welfare performance of sequential pricing mechanisms." *International Economic Review*, Vol 46 (2005), pp. 1321-1360.

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

Park, W. "Mobile phone addiction". *Mobile Communications*, (2005), pp. 253-272.

Pollak, R. A. "The intertemporal cost of living index". *In Annals of Economic and Social Measurement*, Vol. 4 (1975), pp. 179-198.

Sandroni, A. and F. Squintani. "Overconfidence and adverse selection: The case of insurance". Technical report, working paper, 2010.

Spinnewijn, J. "Insurance and perceptions: how to screen optimists and pessimists.*The Economic Journal*, Vol 123 (2013), pp. 606-633.

9 Appendix A

Proof of Lemma 2

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

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

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

The optimization problem of the firm when the consumer is naive hafit forming is:

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

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.

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

Moreover, Δ is the difference between the perceived and the optimal utility of the consumer.

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

We simplify $S(\mathbf{p^N})$ and Δ and delete $\int_{v_{1N}^*}^1 \int_{p_2-\beta}^1 (v_2 + \beta)f(v_2)dv_2 dF(v_1)$ from both.

Then the first order conditions with respect to p_1 are:

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

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

Then the first order condition is:

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

Then

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

Thus,

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

Moreover, the first order conditions with respect to p_2:

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

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

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

Substituting for the optimal threshold and after some algebra [15]

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

Then:

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

[15] Let $p_1 < p_2$

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

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

Substituting (5) and rearranging:

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

Moreover, let for simplicity

$$
\begin{aligned}
A &= 1 - F(p_2 - \beta) + \left(\frac{\int_{v_{1N}^*}^{1} f(v_1) f(p_2 - \beta) dv_1}{f(v_{1N}^*)(1 - F(p_2))} \right) - \frac{F(p_1)^2 f(v_{1N}^*)(1 - F(p_2 - \beta))}{F(p_1)^2 f(v_{1N}^*) + F(v_{1N}^*) f(p_1)} \\
&= \left(\frac{\int_{v_{1N}^*}^{1} f(v_1) f(p_2 - \beta) dv_1}{f(v_{1N}^*)(1 - F(p_2))} \right) + (1 - F(p_2 - \beta)) \left(1 - \frac{F(p_1)^2 f(v_{1N}^*)}{F(p_1)^2 f(v_{1N}^*) + F(v_{1N}^*) f(p_1)} \right) \\
&= \left(\frac{\int_{v_{1N}^*}^{1} f(v_1) f(p_2 - \beta) dv_1}{f(v_{1N}^*)(1 - F(p_2))} \right) + (1 - F(p_2 - \beta)) \left(\frac{F(p_1)^2 f(v_{1N}^*) + F(v_{1N}^*) f(p_1)}{F(p_1)^2 f(v_{1N}^*) + F(v_{1N}^*) f(p_1)} \right) \\
&\quad + (1 - F(p_2 - \beta)) \left(\frac{-F(p_1)^2 f(v_{1N}^*)}{F(p_1)^2 f(v_{1N}^*) + F(v_{1N}^*) f(p_1)} \right) \\
&= \frac{\int_{v_{1N}^*}^{1} f(v_1) f(p_2 - \beta) dv_1}{f(v_{1N}^*)(1 - F(p_2))} + \frac{(1 - F(p_2 - \beta)) F(v_{1N}^*) f(p_1)}{F(p_1)^2 f(v_{1N}^*) + F(v_{1N}^*) f(p_1)} > 0
\end{aligned}
$$

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:

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

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

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

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

Proof of Lemma 3

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 ϵ would preserve the IR_N, would not affect the IC_S and IC_N, and raise profits which contradicts 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. Moreover, 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, thus a contradiction.

Moreover, as we expect the incentive compatibility constraint of the sophisticated consumer binds at the optimum. It could be written as:

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

Proof of Proposition 3

The profits of Naive consumers are:

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

The profits of the Sophisticated consumers are:

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

where

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

because the IC_S and IR_N are binding.

Thus, the profit function for the screening model is:

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

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

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

$$
\begin{aligned}
\frac{d\Pi}{dp_1} &= \frac{d\Pi}{dv_{1N}^*}\frac{dv_{1N}^*}{dp_1} + \frac{d\Pi}{dp_1} = \\
&= f(v_{1N}^*)F(p_1)\Bigg((\gamma-1)(c-p_2)F(p_2-\beta) - (\gamma-1)\left(\int_{p_1}^{1}(v_2-c)f(v_2)dv_2\right) \\
&- 2c(\gamma-1) + \gamma\int_{p_1}^{1}(1-F(v_2))dv_2 - \int_{p_2}^{1}(1-F(v_2))dv_2 + \gamma(p_1+p_2) - p_2 - v_{1N}^*\Bigg) \\
&+ (1-\gamma)(c-p_1)f(p_1)F(v_{1N}^*) \\
&+ \gamma\left((F(p_1)-1)(F(v_{1N}^*) - F(v_{1S}^*)) + \int_{v_{1S}^*}^{v_{1N}^*} f(v_1)dv_1\right) = \\
&= f(v_{1N}^*)F(p_1)\Bigg((\gamma-1)(c-p_2)F(p_2-\beta) - (\gamma-1)\left(\int_{p_1}^{1}(v_2-c)f(v_2)dv_2\right) \\
&- 2c(\gamma-1) + \gamma\int_{p_1}^{1}(1-F(v_2))dv_2 - \int_{p_2}^{1}(1-F(v_2))dv_2 + \gamma(p_1+p_2) - p_2 - v_{1N}^*\Bigg) \\
&+ (1-\gamma)(c-p_1)f(p_1)F(v_{1N}^*) \\
&+ \gamma\left((F(p_1)-1)(F(v_{1N}^*) - F(v_{1S}^*)) + (F(v_{1N}^*) - F(v_{1S}^*))\right)
\end{aligned}
$$

after some algebra[16] and substituting v_{1N}^*:

$$
\begin{aligned}
\frac{d\Pi}{dp_1} = F(p_1)(1-\gamma)f(v_{1N}^*)\Big((p_2-c)F(p_2-\beta) + F(p_1)(c-p_1) + c - p_2\Big) \\
+ f(p_1)(1-\gamma)(c-p_1)F(v_{1N}^*) + F(p_1)\gamma(F(v_{1N}^*) - F(v_{1S}^*)) = 0
\end{aligned}
$$

The first order condition with respect to p_2 is:

16

$$
\begin{aligned}
&(1-\gamma)\int_{p_1}^{1}(v_2-c)f(v_2)dv_2 + \gamma\int_{p_1}^{1}(1-F(v_2))dv_2 - \int_{p_2}^{1}(1-F(v_2))dv_2 \\
&= -(1-\gamma)(1-F(p_1))c + (1-\gamma)(1-p_1F(p_1)) + \gamma(1-p_1) - (1-p_2) - \int_{p_1}^{p_2} F(v_2)dv_2
\end{aligned}
$$

$$
\begin{aligned}
\frac{d\Pi}{dp_2} &= \frac{d\Pi}{dv_{1N}^*}\frac{dv_{1N}^*}{dp_2} + \frac{d\Pi}{dp_2} = \\
&= f(v_{1N}^*)(1-F(p_2))\Bigg((\gamma-1)(c-p_2)F(p_2-\beta) + (1-\gamma)\left(\int_{p_1}^{1}(v_2-c)f(v_2)dv_2\right) \\
&- 2c(\gamma-1) + \gamma\int_{p_1}^{1}(1-F(v_2))dv_2 - \int_{p_2}^{1}(1-F(v_2))dv_2 + \gamma(p_1+p_2) - p_2 - v_{1N}^*\Bigg) \\
&+ (\gamma-1)(c-p_2)(F(v_{1N}^*)-1)f(p_2-\beta) + F(p_2-\beta)((1-\gamma)F(v_{1N}^*) + \gamma F(v_{1S}^*) - 1) = 0
\end{aligned}
$$

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

$$
\begin{aligned}
\frac{d\Pi}{dp_2} &= (F(p_2)-1)(\gamma-1)f(v_{1N}^*)((p_2-c)F(p_2-\beta) + F(p_1)(c-p_1) + c - p_2) \\
&+ (\gamma-1)(c-p_2)(F(v_{1N}^*)-1)f(p_2-\beta) + F(p_2-\beta)(\gamma F(v_{1N}^*) + \gamma F(v_{1S}^*) + F(v_{1N}^*) - 1) \\
&+ \gamma(F(v_{1N}^*) - F(v_{1S}^*)) - F(p_2)(F(v_{1N}^*)-1) = 0
\end{aligned}
$$

Solving the system of the first order conditions:

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

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

Moreover, $p_1 < c$ if the difference of the fractions is negative namely:

$$\gamma f(p_2-\beta)(1-F(v^*_{1N}))(F(v^*_{1N})-F(v^*_{1S}))$$
$$-f(v^*_{1N})(F(p_2)-F(p_2-\beta))(1-F(p_2-\beta))(1-(1-\gamma)F(v^*_{1N})-\gamma F(v^*_{1S}))<0$$

thus when the fraction of sophisticated consumers γ is relatively small:

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

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 $p_2>0$.

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

$$
\begin{aligned}
S^N(\mathbf{p^N}) &= \int_{v_{1N}^*}^1 \big(v_1 - c + \int_{p_{2N}-\beta}^1 (v_2+\beta-c)dv_2\big)dv_1 + v_{1N}^* \int_{p_{1N}}^1 (v_2-c)dv_2 \\
U^S(\mathbf{p^N}) &= \int_{v_{1S}^*}^1 \big(v_1 - p_{1N} + \int_{p_{2N}-\beta}^1 (v_2+\beta-p_{2N})dv_2\big)dv_1 + v_{1S}^* \int_{p_{1N}}^1 (v_2-p_{1N})dv_2 - F^N \\
U^N(\mathbf{p^N})) &= \int_{v_{1N}^*}^1 \big(v_1 - p_{1N} + \int_{p_{2N}}^1 (v_2-p_{2N})dv_2\big)dv_1 + v_{1N}^* \int_{p_{1N}}^1 (v_2-p_{1N})dv_2 - F^N \\
\Delta &= U^N(\mathbf{p^N}) - \tilde{U}(\mathbf{p^N}) \\
\tilde{U}(\mathbf{p^N}) &= \int_{v_{1N}^*}^1 \big(v_1 - p_{1N} + \int_{p_{2N}-\beta}^1 (v_2+\beta-p_{2N})dv_2\big)dv_1 + v_{1N}^* \int_{p_{1N}}^1 (v_2-p_{1N})dv_2 - F^N \\
v_{1N}^* &= p1 + \int_{p_{1N}}^{p_{2N}} (1-v_2)dv_2 \\
v_{1S}^* &= p1 + \int_{p_{1N}}^{p_{2N}-\beta} (1-v_2)dv_2
\end{aligned}
$$

Making the calculation we have:

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

The derivative with respect to p_1:

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

and with respect to p_2:

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

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:

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

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:

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

,where

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

Moreover, 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 and substituting the prices it is reasonable expect for the parameter levels that there is not real root.

10 Appendix B: Partially Naive Consumers

Let a partially naive consumer with habit formation coefficient β and $0 < \theta < 1$. The consumer values at period t is:

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

The optimization problem of the firm is:

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

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.

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

Moreover, Δ is the difference between the true and the perceived of the consumer utility.

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

Thus, Δ 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:

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

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

Then, the first order condition is:

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

Thus,

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

The first order condition with respect to p_2 is:

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

then the respective derivatives are:

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

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

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

[17]

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

Finally solving the above system of equations we get:

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

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

Thus, we see that $p_1 < c$ and $p_2 > c$.

---

[17]

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

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