A new theoretical explanation for three part tariffs - PRELIMINARY -

Pedro Alves

May 8, 2015

Abstract

The model developed here builds upon the standard sequential screening models where consumers do not know their consumption type when signing a contract with a firm for the provision a good. I extend this model by assuming that this interaction happens twice but the consumer, in the second, has an imperfect memory of the first period interaction and his type.

Imperfect consumer memory, as modelled here, is: imperfect when it comes to the values and motivation for past consumption; but perfect when it comes to the prices paid in the past. I show that when this is the case the firm is going to have an incentive to offer contracts commonly called three part tariffs (3PT).

1 Introduction

I formalise in a theoretical model a new argument for why firms choose to offer contracts commonly called three part tariffs (3PT). These contracts are commonly offered by cellphone and internet data companies. For instance, the typical cellphone contract is a 3PT. As the name suggests, this contract has three components: a fixed fee paid before-hand; a number of minutes offered for free (called an allowance) and a price per minute charged when consumption level exceeds the allowance.

Why firms use these contracts is puzzling because the marginal costs in these industries are low but positive. Therefore, when the firms offer 3PTs they are subsidising low units of consumption and overcharging (charging above the

marginal cost) the consumption of higher units. It is not obvious why firms want to do this.

In this paper, I suggest a novel argument for why firms in these industries offer 3PTs. The argument relies on three important observation in these industries. First, the consumers in these industries seem to be only partially informed about their own demand profile. This is natural since the services provided by the firms are relatively hard to track by the consumers. Second, the consumers seem to learn about their demand profile over time. Whether consumers learn from past data is a question that divides economists. However, recent empirical papers (Miravete and Palacios-Huerta, 2014 and Osborne and Grubb, 2010) provide evidence that consumer learning does occur in cellphone markets. Third, these same papers also observe that consumer learning is faster when the consumers have more feedback from their past choices through bills. This suggests that consumers can remember past information more easily when it is expressed as prices instead of just levels of consumption.

I show that, when learning occurs in this way, 3PT contracts maximise the firm's profits. When the consumer's learning comes mainly from inference from the prices they were charged in the past, the firms will have an incentive to distort prices. This creates a strong incentive for the firms to offer contracts with "flat pricing" - that provide less information to the consumers. Finally, I show that the incentives to prevent learning differ over the consumer types the firm is facing which leads to 3PT contracts.

2 Set-up

In every period $t \in \{1, 2\}$, a continuum of consumers contract with one firm over the provision of a non-durable good $q_t \in [0, \infty)$. Each consumer has a demand type $\theta \in \{\theta_1, \theta_2, \theta_3, \theta_4\} \equiv \Theta$ and $\theta_1 < \theta_2 < \theta_3 < \theta_4$. In the first period, the type of each consumer is unknown to both the consumer and the firm. The type of each consumer is constant over time and it is drawn, initially, from an uniform distribution over Θ. All these facts are common knowledge to the firm and consumers.

Timing. Each period $t \in \{1, 2\}$ is composed of three stages stages. First, in the *"recall stage"* each consumer observes his memory m_{t-1} from the past period and decides on a memory function for the current period. Second, in the *"contracting stage"* the firm offers a menu of contracts $\{q_t(\theta, m), P_t(\theta, m)\}$. Each different contract (designed for a consumer with memory m) describes a quantity and payment schedule intended for each realisation of the demand type θ. At this point, each consumer chooses one of the contracts from the menu or rejects them all. When rejecting the contract the consumer gets his outside option normalised to zero. Finally, in the *"consumption stage"* the consumer observes his demand type θ. Then, conditional on the contract accepted, chooses quantity $q_t(\theta', m)$ and pays $P_t(\theta', m)$.

[Insert figure here!]

Consumers and firm. The objective functions of the consumers and firm used here are standard in screening problems. Consumers' utility U is given by the utility of consumption $V(q(\theta'), \theta)$ minus the payment to the firm. For concreteness, we are going to use a particular utility function: $V(q(\theta), \theta) = 2\theta q$. For convenience of computations, assume that consecutive types are equidistant and the common distance between types is $\Delta \equiv \theta_j - \theta_{j-1}$.

The firm maximises expected profits given by the revenues minus convex production costs $C(q)$. To get closed form solutions, assume that $C(q) = q^2$. Assume that $\theta_1 > 6\Delta$, such that the firm serves all types.

Finally, both the consumers and the firm are forward looking. That is, both maximize the sum of payoffs over both periods. However, for the purpose of this draft, I will treat the consumer's decision in the contracting and consumption stage as myopic and the memory decision as forward looking. This is arguably

without loss of generality, but more work needs to be done to show this. [1]

Partial commitment assumption. The firm can only commit to prices within each period. That is, for unmodelled reasons, in the first period the firm cannot commit to prices for transactions in the second period. This means that the contracts are negotiated every period, as described in the timeline above. This is important since any problems with future (consumer) learning could be assumed away if the firm could contract today over future periods. Further, as it is standard in the screening literature, the firm cannot offer exclusive contracts to some consumers. This means that, even if the firm observes the consumer's type, it cannot offer contracts contingent on that information.

Description of memory/learning with imperfect recall. The assumptions on consumer memory are the novel aspect of this model. The consumers in this model are assumed to have imperfect recall. As described in the timing above, all the consumers start uninformed (about their type) in the beginning of the first period. Then, at the end of the first period, each consumer observes his type. However, at the beginning of the second period they might not remember their previously observed type. Instead, they start the second period with a memory of their type that might not be perfect.

Call $m_i(\theta)$ the "memory" of a consumer i in the second period. This memory is relevant since it is used to decide which contract to sign in the second period. Formally, define $m_i(\theta)$ as the set of types the consumer believes he can belong to after observing the (limited) information he recalls from the first period. As discussed below, this will be a sufficient statistic for his type posterior.

A consumer with *perfect* recall would always remember his type's realisation from the first period, such that $m_i(\theta) = \theta$. Here, in the second period, each consumer i will only have access to an imperfect memory m_i which aggregates two memory sources: active memory (m_i^A) and passive memory (m^P). A consumer i will remember his type, such that $m_i(\theta) = \theta$, whenever at least one element of his memory includes this information. On the other hand, a consumer i will be

[1] The argument could follows from these facts: 1). The contract choice of a single consumer is not going to influence the contracts offered by the firm. This follows from the fact that there are a continuum of consumers. 2). In the (second) contracting stage of the first period, the consumers are homogeneous meaning that the firm offers a single contract. Therefore, the potential incentive of consumers to choose the "wrong" contract to learn more about their type are absent. 3) I might need to assume that decision in (third) consumption stage cannot be used to signal information about type.

as uninformed as in the beginning first period when his memory is $m_i(\theta) = \Theta$.

Passive memory. This is a costless memory technology that all consumers have access to. When using the passive memory, the consumers remember the contract they were offered in the first period ($P_1(\theta, m)$) and their first period's realised payment.[2] With this memory, each consumer learns more about his type by comparing his bill with the amount that would have been paid by each type.[3] Therefore, using the passive memory, all the consumers are going to learn their type when the pricing is strictly monotone for all types (i.e., $P'(\theta) > 0$ for all θ). On the other hand, when the first period pricing is the same for all the types (such that $P'(\theta) = 0$ for all θ) nothing is learned about the type using the passive memory.

Formally, define $m^P(\hat{\theta})$ as the set of types that the consumer believes he can belong to using passive memory.[4] Given the assumptions above, this set for a consumer of type $\hat{\theta}$ is given by:

$$m^P(\hat{\theta}) = \{\theta \in \Theta : P_1(\theta) = P_1(\hat{\theta})\}$$

Active memory. This component of memory is modeled as a costly memory technology. In the beginning of each period, the consumer chooses a memory function that maps the type the consumer will observe in the last stage to a choice of "recall" or "not recall". Suppose that the rule chosen by a consumer i of type $\hat{\theta}$ is "recall" when $\theta = \hat{\theta}$. Then $m^A(\hat{\theta}) = \hat{\theta}$ and the consumer pays a cognitive cost F. An alternative rule that chooses "not recall" when $\theta = \hat{\theta}$ would generate $m^A(\hat{\theta}) = \Theta$, meaning that nothing is learned.[5]

[2] This is based on empirical evidence that shows that past prices (and bills) are objects that the consumers remember more easily than other objects.

[3] This could be modelled as a dual self model where the two selves do not share information. Call the consumer's self who does the decision in the first stage and in the second stage, respectively, self 1 and 2. Then if self 1 has a perfect memory and learns in a Bayesian way, but does not share information with self 2 we would get the assumption above.

[4] Given that the priors are uniform and the information from passive learning is not noisy the posterior using passive learning is also uniformly distributed.

[5] The rule makes sure that there are not consistency problems and that the inference when observing $m^A(\theta) = \Theta$ is clear. Alternatively, we could just assume that the consumer, after observing his type, chooses in the last stage whether to remember or not. But this would allow the consumer to manipulate his own memory and always remember his type in equilibrium and never pay the cost.

The consumer will choose to remember a particular type θ' in the next period if the benefit of being informed about his type exceeds the cost F. This is going to happen, as elaborated below, when the second period equilibrium contract (anticipated by the consumer) gives this type θ' a (negative) surplus which is smaller than the negative cost of recall $-F$. If the consumer pays the cost to remember it can reject the second period contract and get his zero outside option instead.

These two memory sources are aggregated into $m_i(\theta)$. Note that memory is an object that can differ across individuals since the active memory is an individual decision, contrary to the passive memory which is common to all consumers.

The rule cannot be decided at the point at which the consumer knows θ. The reason for this is that observing $m^A(\theta) = \Theta$ by itself is informative. This can be easily proved by contradiction. Suppose that the consumer could decide on a rule ex post. Then, the consumer that observes his type to be θ' could always define the rule $m^A(\theta) = \emptyset$ when $\theta = \theta'$ and $m^A(\theta) = \theta$ otherwise. This would mean that the consumer would never pay F and would always remember his type.

3 The second period's problem

This section shows how the second period problem of the firm can be solved and establishes two important facts. First, consumer learning (from the first period) is going to reduce the profits of the firm in the second period. Second, the information/memory environment where the second period (optimal) contract generates a consumer type with the lowest surplus is the one where all types are uninformed. This is going to play a role in the active memory decision analysed in section 5.

The fact that this is the last period makes both the firm and the consumer's problem simpler. The consumer will never pay to remember his type so the memory decision is mute. Similarly, the optimal contract offered in the second period maximises the own period's profits only since the firm cannot commit to any second period contract in the first period. Finally, the memory technology generates information types that form partitions of the type space. This means that this model can be seen as having distributions of types that FOSD each other. This means that the problem can be analysed as a standard sequential screening problem with tools developed by Courty and Li (2001).

3.1 Formalising the problem

The firm is going to offer a menu of contracts. Each contract is intended for a consumer with memory m and it is written as $\{P(\theta, m), q(\theta, m)\}$. Define the set of all memories generated by the pricing in the first period P_1 as $M(P_1)$. These sets form a partition of the type space.

It is going to be useful to define some notation. The utility of a consumer of type θ who, in the third stage of any period, chooses a quantity designed for type θ' (in contract of memory m) is $U(\theta, \theta'|m) \equiv V(q(\theta'), \theta)$. The utility of a truthful consumer is $U(\theta|m) \equiv U(\theta, \theta|m)$. The expected utility of a consumer, in the first stage, with memory m who chooses contract designed for consumer with memory m' is $E[U(\theta, m')|m]$. Then, the firm's problem can be written as:

Therefore, the firm's problem in the second period can be written as:

Problem 1.

$$Max_{q(\theta,m)\geq 0\ ,\ P(\theta,m)}\ E\left[P(\theta, m) - C(q(\theta, m))\right]$$

subject to:

(for all information sets $m \in M(P_1)$)

$$\textbf{\textit{Incentive compatibility (IC)}}\ \ U(\theta,\theta|m) \geq U(\theta,\theta'|m) \ \ , \ \ \forall\theta,\theta' \in m$$

$$\textbf{\textit{Global IC}}\ \ E[U(\theta,m)|m] \geq E[U(\theta,m')|m] \ \ , \ \ \forall m' \in M(P_1)$$

$$\textbf{\textit{Participation constraint}}\ \ E[U(\theta,m)|m] \geq 0 \ \ , \ \ \forall m \in M(P_1)$$

The *incentive compatibility* constrains each consumer to be truthful within the contract that he chose. In other words, this constraint, when satisfied, means that the consumers report their types truthfully in the second stage when they observe their types. The *global incentive compatibility* imposes that consumers who observed m prefer contracts designed for them than other contracts intended to consumers who observed some other information m'. When nothing is learned in the first period, this constraint does not exist since the firm offers only one contract. Finally, the *participation constraint* imposes that each contract for a memory m offers a non-negative expected surplus to consumers.

3.2 Main features of the contracts in the second period

Define the partition of the type space generated by the memories of each consumer $m(\theta)$ in a symmetric equilibrium where all consumers use active memory in the same way (such that we can drop the dependence of memory on individual consumer)[6]:

Definition 1. $\tilde{m}(P^1)$ *is the partition of the type space composed of memories of consumers of different types generated by some first period pricing P^1.*

I define $\Pi^2_{\tilde{m}}$ as period 2's profits as a function of memory partition $\tilde{m}$. For instance, when $\tilde{m} = \{\{1,2\}\{3,4\}\}$ second period profits are written as $\Pi^2_{12|34}$.

[6]Below it is shown that this is going to be the only equilibrium

Proposition 1. *Learning decreases profits: memories with finer partitions generate lower second period profits:*

$\Pi^2_{1234} > \Pi^2_{123|4} > \Pi^2_{12|3|4} > \Pi^2_{1|2|3|4}$

Proof:

1) The PC, IC and global IC differ depending on the memory partition that the firm faces. Define two memory partitions $\tilde{m}_1$ and $\tilde{m}_2$. And assume that $\tilde{m}_1$ defines a (memory) partition that is finer than $\tilde{m}_2$.

Then, the set of contracts that are feasible under the constraints defined by $\tilde{m}_1$ are contained in the set of contracts that satisfy constraints defined by $\tilde{m}_2$.

The reason for this is that PCs of $\tilde{m}_2$ imply the PCS of $\tilde{m}_1$. The same is true for the global ICs and IC of $\tilde{m}_2$. Hence, the profits are weakly larger when the firm faces constraints given by $\tilde{m}_2$.

2) For the details of the computation of profits in the relevant memories[7] see appendix 1.

Finally, use the previously defined notation $U(\theta_i|m)$ for the surplus of a consumer with type θ in contract m. Different histories generate contracts with different surpluses. Define the lowest surplus (over all consumer types) in a second period contract with memory partition $\tilde{m}$ as $\underline{U}_{\tilde{m}}$.[8] An important fact in the surpluses generated by second period contracts is that:

Proposition 2. *The minimum surpluses of contracts with different memories have the following relation:* $\underline{U}_{1234} < \underline{U}_{123|4} < \underline{U}_{12|34} < \underline{U}_{1|234} \; = \; 0$

Proof:

1) Direct computation gives us these expressions (as seen in appendix 1):

$$\underline{U}_{1234} \;=\; -\frac{3}{4}(V_2(q_1) - V_1(q_1)) - \frac{1}{2}(V_3(q_2) - V_2(q_2)) - \frac{1}{4}(V_4(q_3) - V_3(q_3))$$

$$\underline{U}_{123|4} \;=\; -\frac{2}{3}(V_2(q_1) - V_1(q_1)) - \frac{1}{3}(V_3(q_2) - V_2(q_2))$$

$$\underline{U}_{12|3|4} \;=\; -\frac{V_2(q_1) - V_1(q_1)}{2}$$

[7] This excludes memories that could only have been generated by decreasing prices, since this case is excluded by the monotonicity constraint.

[8] This is $min_{\;\theta\in\Theta}\; U(\theta|m)$ for surpluses in contracts with history $\tilde{m}$

2) Use the fact that quantities are distorted more in contracts with more learning. **(this needs to be improved)**

(I need a general proof of this fact!)

4 First period's problem

4.1 Formalising the problem

By assumption, all the consumers start uninformed in the first period. The firm offers a single contract to the homogeneous consumers that it faces in the first period. Then, if the firm was myopic - i.e., only maximised the first period profits - the firm's problem could be written as:

Problem 2.

$$Max_{q(\theta),P(\theta)} \;\; E[P(\theta) - C\big(q(\theta)\big)]$$

subject to:

Incentive compability $U(\theta, \theta|m) \geq U(\theta, \theta'|m)$
and

$$\textbf{\textit{Participation constraint}} \;\; E[U(\theta, m)|m] \geq 0$$

and would therefore offer welfare maximising contracts.

Instead, the forward looking firm in this model will choose to distort the first period's contract to "prevent learning of some types" in the second period.[9] I will say that the firm *prevents learning of types* $\theta \in \Theta^P \subset \Theta$ when $m(\theta) = \Theta^P$ for all $\theta \in \Theta^P$. And more generally I will say that the firm chooses a *pricing which prevents consumer learning* when the firm is *preventing learning of some types* as defined above.

Nevertheless, in the first period, there is no asymmetric information so that the firm does not have to pay any information rents and can extract the full consumer surplus by making the PC bind. Suppose for instance that the firm wants to optimally prevent learning of the types $\theta \in \Theta^P \subset \Theta$, then the first period's problem can be rewritten as:

[9]Note that in some cases this might not be possible to do with pricing since the consumer might use active memory to remember his type.

Problem 3.

$$Max_{q(\theta),P(\theta)} \;\; E[V\big(q(\theta),\theta\big) - C\big(q(\theta)\big)]$$

subject to:

Incentive compability $U(\theta,\theta|m) \geq U(\theta,\theta'|m)$

and

Preventing Learning constraint $P(\theta_j) = P(\theta_k) \quad \forall\theta_j \ , \ \theta_k \ \in \Theta^P$

Define the profits in the first period of contracts that generate passive memory partitions $\tilde{m}^P$ in the second period as $\Pi^1_{\tilde{m}^P}$, then the following is going to be true:

Proposition 3. *Preventing learning is costly in terms of first period's profits. In particular,*

$\Pi^1_{1234} < \Pi^1_{123|4} < \Pi^1_{12|3|4} < \Pi^1_{1|2|3|4}$

Proof:

1) Using the incentive compatibility, you can see that if $P(\theta_j) = P(\theta_k)$ then $q(\theta_j) = q(\theta_k)$. This follows from the fact that the consumers always have positive marginal utility from consumption.

2) Therefore, the more types the firm wants to prevent learning the more constraints it faces on the quantities that it can implement. See appendix 2 for more details.

(unnecessary for now...)

3) In more detail. Define $W^*(\theta)$ as the welfare of an efficient allocation offered to type θ. Define W^* as the sum of welfare for all types.

The welfare of efficient types is convex but convexity is decreasing. That is, $\frac{\partial^3 W^*(\theta)}{\partial \theta^3}$

4) When the firm does not prevent learning the profits are just given by the sum of total welfare, that is W^*. This comes from the fact that there is no asymmetric information.

When the firm prevents learning of a set (interval) of types $\theta \in \Theta^P$ the profits are given by $prob(\theta \in \Theta^p)W^*(\mu) \; + \; (1 - prob(\theta \in \Theta^p))W^*(\theta)$ (notation is not great in discrete case...)

5) Decreasing convexity implies that profits are smaller than efficient ones.

5 Optimal equilibrium contracts (without active memory)

Definition 2. *The total profits of a dynamic contract, which uses a pricing in the first period that generates (through passive memory) the memory partition* $\tilde{m}$ *in the second period, are* $\Pi_{\tilde{m}}$*. Such that* $\Pi_{\tilde{m}} \equiv \Pi^1_{\tilde{m}} + \Pi^2_{\tilde{m}}$.

For example when $\tilde{m} = \{\ \{\theta_1, \theta_2, \theta_3\}, \{\theta_4\}\ \}$ this is going to be written as $\Pi_{123|4}$.

Types of contracts. I differentiate the contracts by their first period. I call 3PT to contracts that prevent learning of low type consumers and not high types. These have profits given by $\Pi_{123|4}$ or $\Pi_{12|3|4}$. I call reversed-3PT to contracts that prevent learning of high type consumers and not low types.These have profits given by $\Pi_{1|234}$ or $\Pi_{1|2|34}$. I call the contract that does not prevent learning of any types the myopic contract. This has profits given by Π_{1234}. I call the contract that prevents learning of all types a pooling contract. This has profits given by $\Pi_{1|2|3|4}$. I call the contract that prevents learning of high and low types separately a double pooling contract. This has profits given by $\Pi_{12|34}$.

It is useful to establish relationship between profits of these contracts. First it can be seen that between the most extreme contracts: pooling and myopic contract the firm is always going to choose pooling.

Lemma 1. *Profits of pooling contracts are higher than profits of screening (myopic contract):*

$$\Pi_{1234} \ > \ \Pi_{1|2|3|4}$$

Proof:

0) The profits of the myopic contract in the first period are equal to the profits of the pooling contract in the second period. Therefore, the total dynamic profits of these two contracts can be compared by evaluating the profits of myopic contract in the second period and pooling contract of the first period.

1) Profits of pooling are equal to the efficient welfare of the average type (i.e., $E(\theta) = \mu$) which is W^*_μ. Use this fact and write the profits of pooling as

the difference from the fully efficient profits to get this:

$$W^* - \frac{\partial W}{\partial \theta}(\mu)(\theta_4+\theta_1-2\mu) - \int_{\mu}^{\theta_4} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta_4-\theta)\ \mathrm{d}\theta - \int_{\theta_1}^{\mu} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta-\theta_1)\ \mathrm{d}\theta$$

Where the first term is zero by symmetry (of uniform distribution).

2) That is profits of screening are equal to profits of contract minus information rents. These can also be written as the efficient welfare of lower types. In particular, each type is decreased such that new types are defined as: $\theta_i^{screen} = \theta_i - \frac{prob(\theta>\theta_i)}{prob(\theta=\theta_i)}\Delta$. This means that profits of myopic contract in the second period (which are the textbook's "screening profits") can be written as:

$$\frac{1}{4}(\ W^*_{\theta_4} + W^*_{\theta_3-\Delta} + W^*_{\theta_2-2\Delta} + W^*_{\theta_1-3\Delta}\ )$$

which can be rewritten as:

$$W^* - \ \frac{1}{4}\ (\ \frac{\partial W}{\partial \theta}(\theta_1)3\Delta + \frac{\partial W}{\partial \theta}(\theta_2)2\Delta + \frac{\partial W}{\partial \theta}(\theta_3)\Delta\ )$$

$$-\ \frac{1}{4}(\int_{\theta_1-3\Delta}^{\theta_1} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta_1-\theta)\ \mathrm{d}\theta + \int_{\theta_2-2\Delta}^{\theta_2} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta_2-\theta)\ \mathrm{d}\theta + \int_{\theta_3-\Delta}^{\theta_3} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta_3-\theta)\ \mathrm{d}\theta)$$

3) Convexity is weakly decreasing (i.e., $\frac{\partial^3 W^*(\theta)}{\partial \theta^3} \leq 0$)

Weakly decreasing convexity is true for $V(q;\theta) = \theta v(q)$ when $v()$ is concave and $c()$ is convex.

4) Using the previous point on the equations above it is easy to show that the profits of pooling are always larger than profits of screening.

Intuition Pooling and screening contracts both impose distortions compared to a fully efficient contract. The distortions of screening coming from information rents are considerably larger. However, the distortions in pooling contracts are going to affect all types while the screening distortions are going to be concentrated mainly on the lower types. Decreasing convexity guarantees that the distortions for higher types however are not growing without bounds.

Lemma 2. *Profits of 3PT are increasing in the amount of types that are prevented from learning:*

$$\Pi_{123|4} \;>\; \Pi_{12|3|4}$$

Proof:

1) 3PTs differ in the first period by the cost of pooling. Longer 3PT have higher costs since they pool more types:

(Short 3PT)

$$W^* \;-\int_{\mu_M}^{\theta_2} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta_2-\theta)\,\mathrm{d}\theta \;-\int_{\theta_1}^{\mu_M} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta-\theta_1)\,\mathrm{d}\theta$$

(Long 3PT)

$$W^* \;-\int_{\mu_L}^{\theta_3} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta_3-\theta)\,\mathrm{d}\theta \;-\int_{\theta_1}^{\mu_L} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta-\theta_1)\,\mathrm{d}\theta$$

Net loses of more pooling in the first period can therefore be written as:

....

2) In the second period, the allocations for the high types who have learned are just screening allocations. The allocations for low types who have not learned are distorted by the fact that they influence the surplus the firm needs to give to higher types:

$$q(\theta_i) \;=\; \theta_i - \frac{prob(\theta \leq \theta_i)}{prob(\theta = \theta_i)} \frac{prob(\theta > \theta^*)}{prob(\theta \leq \theta^*)} \Delta$$

(for types below or equal the boundary type in the pooling section θ^* + for the case of $V = \theta q$)

Or similarly the profits can be written as facing lower types $\tilde{\theta} = \theta_i - \frac{prob(\theta \leq \theta_i)}{prob(\theta = \theta_i)} \frac{prob(\theta > \theta^*)}{prob(\theta \leq \theta^*)} \Delta$.

Which means that the gains (in the second period) of longer 3PT can be written as:

$$\int_{\theta_2 - 2\Delta}^{\theta_2 - \frac{2}{3}\Delta} \frac{\partial W}{\partial \theta}(\theta)\frac{1}{4}\,\mathrm{d}\theta \;+\; \int_{\theta_1 - \Delta}^{\theta_1 - \frac{1}{3}\Delta} \frac{\partial W}{\partial \theta}(\theta)\frac{1}{4}\,\mathrm{d}\theta$$

(*Rewrite this in a more general way that allows to generalise to more types. Check 5 type solution for guidance on pattern...*)

which you can rewrite as:

$$\frac{\partial W}{\partial \theta}(\theta_2 - 2\Delta)\frac{4}{3}\Delta \; + \; \int_{\theta_2 - 2\Delta}^{\theta_2 - \frac{2}{3}\Delta} \frac{\partial^2 W}{\partial \theta^2}(\theta)\; (\theta_2 - \frac{2}{3}\Delta - \theta)\frac{1}{4}\; \mathrm{d}\theta$$

and

$$\frac{\partial W}{\partial \theta}(\theta_1 - \Delta)\frac{2}{3}\Delta \; + \; \int_{\theta_1 - \Delta}^{\theta_1 - \frac{1}{3}\Delta} \frac{\partial^2 W}{\partial \theta^2}(\theta)\; (\theta_1 - \frac{1}{3}\Delta - \theta)\frac{1}{4}\; \mathrm{d}\theta$$

3) Remaining argument relies on showing that gains outweight loses.

Lemma 3. *Profits of pooling are higher than profits of 3PTs*

$$\Pi_{1234} \; > \; \Pi_{123|4}$$

Proof:

1) Follows from previous lemma.

Lemma 4. *Profits of I-3PTs are increasing in the types that are prevented from learning.*

$$\Pi_{1|234} \; > \; \Pi_{1|2|34}$$

Proof:

1) Costs in the first period of longer I-3PT come from more types being pooled. The costs of this are given by:

$$\int_{\mu_L}^{\theta_4} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta_4 - \theta)\; \mathrm{d}\theta \; + \int_{\theta_2}^{\mu_L} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta - \theta_2)\; \mathrm{d}\theta$$

minus

$$\int_{\mu_M}^{\theta_4} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta_4 - \theta)\; \mathrm{d}\theta \; + \int_{\theta_3}^{\mu_M} \frac{\partial^2 W}{\partial \theta^2}(\theta)(\theta - \theta_3)\; \mathrm{d}\theta$$

2) The allocation in the second period of I-3PT are:[10]

$$q(\theta^*) = \theta^* \; - \frac{Prob(\theta > \theta^*)}{Prob(\theta = \theta^*)}\Delta \; - \sum_{\theta=\theta^*+\Delta}^{\theta_4} (\theta - \theta^*)$$

(where θ^* is the last type prevented from learning)

$$q(\theta_i) = \theta_i \; - \frac{Prob(\theta > \theta_i)}{Prob(\theta = \theta_i)}\Delta \; = q^{screen}(\theta_i)$$

for $\theta < \theta^*$

3) Consider, the second period benefits of more types being prevented to learn when threshold moves from θ^* to $\theta^* - \Delta$.

This leads to a **benefit** from profits from θ^* given by:

$$W^*(\theta^*) - W^*(\; \theta^* - z(\theta^*) - \sum_{\theta=\theta^*+\Delta}^{\theta_4} (\theta - \theta^*) \;)$$

where $z(\theta^*)$ are the usual screening information rents paid to higher types given by $\frac{Prob(\theta > \theta^*)}{Prob(\theta = \theta^*)}\Delta$. This can be rewritten as:

$$\int_{\theta^* - z(\theta^*) - \sum_{\theta=\theta^*+\Delta}^{\theta_4} (\theta - \theta^*)}^{\theta^*} \frac{\partial W^*}{\partial \theta}(\theta)\; \mathrm{d}\theta$$

And a **cost** given by the increased distortion at new threshold:

$$W^*(\; \theta^* - \Delta - z(\theta^* - \Delta) \;) - W^*(\; \theta^* - \Delta - z(\theta^* - \Delta) - \sum_{\theta=\theta^*}^{\theta_4} (\theta - (\theta^* - \Delta)) \;)$$

which can also be rewritten as:

$$\int_{\theta^* - \Delta - z(\theta^* - \Delta) - \sum_{\theta=\theta^*}^{\theta_4} (\theta - (\theta^* - \Delta))}^{\theta^* - \Delta - z(\theta^* - \Delta)} \frac{\partial W^*}{\partial \theta}(\theta)\; \mathrm{d}\theta$$

Importantly, note that:

[10] Here I am using that $\sum_{\theta=\theta^*+\Delta}^{\theta_4} P_\theta = \sum_{\theta=\theta^*+\Delta}^{\theta_4} V_\theta(q_\theta) - \sum_{\theta=\theta^*+\Delta}^{\theta_4} (V_\theta(q_{\theta^*}) - V_{\theta^*}(q_{\theta^*})) - \sum_{\theta=\theta^*+\Delta}^{\theta_4} (V_{\theta^*}(q_{\theta^*}) - P_{\theta^*})$

$$\sum_{\theta=\theta^*}^{\theta_4}(\theta-(\theta^*-\Delta)) = \sum_{\theta=\theta^*+\Delta}^{\theta_4}(\theta-\theta^*)+z(\theta^*)$$

(Monotonicity constraint is being ignored here! IF it is more likely that it binds for contracts with smaller flat region then this is not a problem...)

4) Remaining argument relies on showing that distortion from more pooling is small compared to the net benefits identified in 3)

Lemma 5. *Profits of 3PT are higher than I-3PT with same lenght flat region:*

$$\Pi_{123|4} \ > \ \Pi_{1|234}$$

and

$$\Pi_{12|3|4} \ > \ \Pi_{1|2|34}$$

Proof:

1)

Lemma 6. *Profits of 3PT are higher than double pooling:*

$$\Pi_{123|4} \ > \ \Pi_{12|34}$$

and

$$\Pi_{12|3|4} \ > \ \Pi_{12|34}$$

Finally, we can find the optimal contract:

Proposition 4. *Pooling contracts are always optimal.*

$$\Pi_{1234} \ > \ \Pi_{\tilde{m}}$$

For $\tilde{m} \neq 1234$

Proof:

1) The profits of pooling contracts are the highest given previous results shown in Lemma 1-6.

This means that it is always going to be optimal to prevent learning of all types in the first period (pooling contract) if these contract can be offered.

6 Optimal recall decision (using active memory)

The memory function decides for which types, observed in the last stage of the first period, the consumer is going to pay the cost F to remember. As mentioned above the existence of a rule decided ex ante is important because the (bayesian) consumer is also going to learn when there is no recall.[11] That is, when $m^A = \Theta$ the consumer is going to know for which types the previous memory function would have chosen not to pay the recall cost.

A few facts are useful to note:

First, the recall decision of an individual consumer will never influence the contract decision of the firm in the second period. This comes from the fact that there is a continuum of consumers. This implies that the recall decision is always made taking as given the (equilibrium) second period's contracts.[12]

Second, every consumer will choose the same memory function. This comes from the fact that the memory function problem is the same for all the consumers. The consumers are homogeneous at the time of choice of memory function and they are going to face the same menu of contracts in the second period.

Proposition 5. *The consumer will use the active memory when:*

1) $F \in (0, -\underline{U}_{12|34})$ and the consumer anticipates that the memory partition in the second period is $\tilde{m} = \{\ \theta_1, \theta_2, \theta_3, \theta_4\}$, $\tilde{m} = \{\ \{\theta_1, \theta_2, \theta_3\}, \{\theta_4\}\ \}$, or $\tilde{m} = \{\ \{\theta_1, \theta_2, \}\{\theta_3\}, \{\theta_4\}\ \}$.

2) $F \in (-\underline{U}_{12|34}, -\underline{U}_{123|4})$ and the consumer anticipates that the memory partition in the second period is $\tilde{m} = \{\ \theta_1, \theta_2, \theta_3, \theta_4\}$ or $\tilde{m} = \{\ \{\theta_1, \theta_2, \theta_3\}, \{\theta_1\}\ \}$.

3) $F \in (-\underline{U}_{123|4}\ ,\ -\underline{U}_{1234})$ and the consumer anticipates that the memory partition in the second period is $\tilde{m} = \{\theta_1, \theta_2, \theta_3, \theta_4\}$.

Proof:

The active memory, in this model, has similar implications to assuming that there is a limited liability constraint on the consumer side in the second period.

In equilibrium, the consumer anticipates the choice of the firm in the second period. Therefore. the consumer benefits from recall only when the type he

[11] A rule decided ex post could be manipulated to avoid costs of recall.

[12] So there is no need to assume anything about the observability of the recall decision by the firm.

observes in the first period is going to have a negative surplus in the second period contract. In this case, the consumer can choose to remember his type and therefore not accept this second period contract. This means getting zero surplus instead of a negative surplus.

This means that if consumers anticipate a memory partition $\tilde{m}$ they will choose to remember their type (for some type realisations) whenever:

$$\underline{U}_{\tilde{m}} < -F$$

That is, when the benefit of recall - not taking a second period contract which has negative surplus - exceeds its cost F.

This also implies that the consumer will never pay to remember in cases where passive memory is informative. Similarly, the consumer will never use the active memory if the cost is higher than the worst possible surplus expected in the second period.

7 Optimal contracts with active memory

Importantly, the profits for the contracts analysed above are computed under the assumption that the consumers do not use the active memory. This is without loss of generality as the following lemma shows:

Lemma 7. *Active memory is never used in equilibrium.*

Proof (intuition):

Suppose there is an equilibrium where the firm offers a contract in each period such that some consumer's types decide to use the active memory. If these consumer's types use the active memory, then the passive memory cannot provide them with information about their types. This can only happen when the firm is pooling over those types. It can be shown that the firm can make a profitable deviation over this contract. Suppose that the firm stops pooling these types in the first period and offers them the efficient quantities instead. Then, the first period's profits will increase and the second period profits will not change since the second period's memory partition will be the same as before.

An important implication of this lemma is that the active memory is going to restrict the set of contracts the firm can choose from. This takes us to our main result:

Result 1. *The optimal contracts in this environment are: -Pooling contracts when $F > -\underline{U}_{1234}$ -Three part tariffs (3PT) when $F \in (-\underline{U}_{123|4}\ ,\ -\underline{U}_{1234})$ -Inverse 3PT when $F \in (0, -\underline{U}_{123|4})$*

Proof (intuition):

1. Use lemma 4, which shows that without active memory (when $F > -\underline{U}_{1234}$) pooling contract is the most profitable one.

$$\Pi_{1234} \ > \ \Pi_{\tilde{m}}$$

2. It is also true that 3PT are going to be more profitable than other contracts (i.e., inversed-3PT and double pooling) whenever[13]:
$\theta_1 > \frac{(38-\frac{3}{2})\Delta}{4}$

3. When $F \in (-\underline{U}_{12|3|4}\ ,\ -\underline{U}_{1234})$ the pooling contract are not be feasible, but 3PTs are. This is a direct consequence of lemma 7.

4. When $F \in (0, -\underline{U}_{12|3|4})$ only inversed 3PTs are going to be feasible.

[13]3PT are always going to be more profitable than inversed-3PT for the same length of types which are prevented to learn.

8 Appendix 1 - seven cases of optimal second period profits

To keep notation simple, I use V_i for $V(., \theta_i)$. Similarly, I use P_i for $P(\theta_i)$ and q_i for $q(\theta_i)$.

8.1 "All types are uninformed" case

When no consumer has learned his type such that $\tilde{m} = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, the firm offers a single contract. The unique participation constraint (PC) is given by:

$$E[U(\theta)] \geq 0$$

Which is equivalent to the PC that the firm faces in the first period. In this case, in the second period, the firm always implements the efficient quantities. That is, $q(\theta_i) \equiv q_i = \theta_i$. The second period profits are:

$$\Pi^M = \frac{(\theta_1)^2 + (\theta_2)^2 + (\theta_3)^2 + (\theta_4)^2}{4}$$

There is a range of prices that implement this allocation. Therefore, the pricing is not explicitly identified. Pricing, however, has to satisfy the following constraints (respectively $IC_{4,3}$, $IC_{3,2}$, $IC_{2,1}$ and PC):

$$V_3(q_4) - V_3(q_3) \leq P_4 - P_3 \leq V_4(q_4) - V_4(q_3)$$

$$V_2(q_3) - V_2(q_2) \leq P_3 - P_2 \leq V_3(q_3) - V_3(q_2)$$

$$V_1(q_2) - V_1(q_1) \leq P_2 - P_1 \leq V_2(q_2) - V_2(q_1)$$

$$P_1 + P_2 + P_3 + P_4 = V_4(q_4) + V_3(q_3) + V_2(q_2) + V_1(q_1)$$

Therefore, the **surplus of the lowest type** can take different values. Define $\underline{U}_{\tilde{m}}$ as the surplus of the type with the lowest surplus in the optimal second period contract when memory partition is $\tilde{m}$. That is, $\underline{U}_{\tilde{m}} = min_{\theta \in \Theta} U(\theta|m)$.

In this case, θ_1 is the type with lowest surplus. The highest possible surplus of θ_1 is:

$$\underline{U}_{1234} = V_1(q_1) - P_1 = V_1(q_1) - \frac{V_1(q_1) - 2V_2(q_2) - V_3(q_3) + 3V_2(q_1) + 2V_3(q_2) + V_4(q_3)}{4}$$

where P_1 is the lowest price that satisfies the constraints above.

Which can be simplified (CHECK):

$$\underline{U}_{1234} = -\frac{3}{4}(V_2(q_1) - V_1(q_1)) - \frac{1}{2}(V_3(q_2) - V_2(q_2)) - \frac{1}{4}(V_4(q_3) - V_3(q_3))$$

8.2 "All types are informed" case

When all consumers have learned their type, such that $\tilde{m} = \{\ \{\theta_1\}, \{\theta_2\}, \{\theta_3\}, \{\theta_4\}\ \}$ there are 4 PCs one for each type. This is the textbook case of screening. In this case, the firm always implements quantities for each type that trade-off the gains for that type and the information rents paid to types above not to deviate to this contract. Quantities are, therefore, given by $q_4 = \theta_4$ $q_3 = 2\theta_3 - \theta_4$, $q_2 = 3\theta_2 - 2\theta_3$ and $q_1 = 4\theta_1 - 3\theta_2$. The second period profits are:

$$\Pi^2_{1|2|3|4} = \frac{(4\theta_1 - 3\theta_2)^2 + (3\theta_2 - 2\theta_3)^2 + (2\theta_3 - \theta_4)^2 + (\theta_4)^2}{4}$$

The **surplus** of the lowest type is zero (by definition). That is, $\underline{U}_{1|2|3|4} = 0$.

8.3 θ_1 and θ_2 are uninformed case

Suppose that, after the first period, the two lowest types are uninformed in the second period. That is, $\tilde{m} = \{\ \{\theta_1, \theta_2\}, \{\theta_3\}, \{\theta_4\}\ \}$. The PCs are defined for the two highest types and the uninformed types. The participation constraint is binding for the lower (uninformed) types but it cannot be binding for higher types. This comes from the fact that when global IC holds for higher types PC is also going to hold. Type θ_3 when deviating to lower contract will always take quantity q_2 designed for θ_2.

Prices for lowest types have to be the highest ones that satisfy the ICs and PC. This means that:

$$P_1 = \frac{V_2(q_2) + V_2(q_1)}{2} - \frac{V_2(q_2) - V_1(q_1)}{2}$$

$$P_2 = \frac{V_2(q_2) + V_2(q_1)}{2} + \frac{V_2(q_2) - V_1(q_1)}{2}$$

Then global IC gives us that $V_3(q_3) - P_3 \geq V_3(q_2) - P_2$
which gives us that:

$$P_3 = V_3(q_3) - (V_3(q_2) - V_2(q_2)) - V_2(q_1) + \frac{V_1(q_1) + V_2(q_1)}{2}$$

and using $IC_{4,3}$

$$P_4 = V_4(q_4) - (V_4(q_3) - V_3(q_3)) - (V_3(q_2) - V_2(q_2)) - V_2(q_1) + \frac{V_1(q_1) + V_2(q_1)}{2}$$

Using these facts, the quantities implemented in the second period are $q_4 = \theta_4$ $q_3 = 2\theta_3 - \theta_4$, $q_2 = 3\theta_2 - 2\theta_3$ and $q_1 = 2\theta_1 - \theta_2$.[14] Therefore, the second period profits are:

$$\Pi^2_{12|3|4} = \frac{(\theta_4)^2 + (2\theta_3 - \theta_4)^2 + (3\theta_2 - 2\theta_3)^2 + (2\theta_1 - \theta_2)^2}{4}$$

Finally, the **surplus** of the lowest type is:

$$\underline{U}_{12|3|4} = V_1(q_1) - \frac{V_2(q_1) + V_1(q_1)}{2}$$

which can be simplified to:

$$\underline{U}_{12|3|4} = -\frac{V_2(q_1) - V_1(q_1)}{2}$$

8.4 θ_3 and θ_4 are uninformed case

In this case, the two highest types are uninformed in the second period such that $\tilde{m} = \{\ \{\theta_1\}, \{\theta_2\}, \{\theta_3, \theta_4\}\ \}$. As before only the PC for the lowest type is binding, that is the one for θ_1. The remaining prices are determined by the global IC from type θ_2 and the global IC from the uninformed types θ_3 and θ_4.

[14]Note that $q_1 = q_2$ when equidistant types.

This last global IC makes sure that the consumer that is uninformed but knows that he is either a consumer of type θ_3 or θ_4 does not want to deviate to the contract of type θ_2.

Following these steps, the firm implements $q_4 = \theta_4$, $q_3 = \theta_3$, $q_2 = 3\theta_2 - \theta_4 - \theta_3$ and $q_1 = 4\theta_1 - 3\theta_2$. The second period profits are:

$$\Pi^2_{1|2|34} \;=\; \frac{(\theta_4)^2 + (\theta_3)^2 + (3\theta_2 - \theta_4 - \theta_3)^2 + (4\theta_1 - 3\theta_2)^2}{4}$$

The surplus of θ_1 is zero. In this case, the prices P_3 and P_4 are not identified by constraints. Using the lowest possible price for θ_3 we can find the surplus of this type:

$$V_3(q_3) - P_3 \;=\; V_3(q_3) - \frac{1}{2}\big[V_4(q_3) + V_3(q_3) - V_4(q_2) - V_3(q_2)\big]$$

Which is always positive.

8.5 θ_2 θ_3 and θ_4 are uninformed case

Here, the three highest types are uninformed in the second period. That is, $\tilde{m} = \{\ \{\theta_1\}, \{\theta_2, \theta_3, \theta_4\}\ \}$. Here, as before the PC is binding for the lowest type and the global IC of higher types is going to identify the prices. That is, the following has to be true:

$$\frac{V_2(q_2) - P_2 + V_3(q_3) - P_3 + V_4(q_4) - P_4}{3} \geq \frac{V_2(q_1) - P_1 + V_3(q_1) - P_1 + V_4(q_1) - P_1}{3}$$

Therefore, the firm implements $q_4 = \theta_4$, $q_3 = \theta_3$, $q_2 = \theta_2$ and $q_1 = 4\theta_1 - \theta_4 - \theta_3 - \theta_2$. The second period profits are:

$$\Pi^2_{1|234} \;=\; \frac{(\theta_4)^2 + (\theta_3)^2 + (\theta_2)^2 + (4\theta_1 - \theta_4 - \theta_3 - \theta_2)^2}{4}$$

The surplus of θ_1 is zero. In this case, the prices P_2, P_3 and P_4 are not identified by constraints. Using the lowest possible price for θ_2 we can find the surplus of this type:

$$\underline{U}_{1|234} \;=\; V_2(q_2) - P_2$$

which is positive.

8.6 θ_1 θ_2 and θ_3 are uninformed case

The three lowest types are uninformed in the second period. That is, $\tilde{m} = \{\ \{\theta_1, \theta_2, \theta_3\}, \{\theta_4\}\ \}$.

The following global IC is going to identify the prices for the highest type:

$$V_4(q_4) - P_4 \geq V_4(q_3) - P_3$$

and the prices for the lowest types have to be the highest possible ones that satisfy the ICs and PC:

$$P_1 = \frac{V_1(q_1) + V_3(q_2) + 2V_2(q_1) - V_2(q_2)}{3}$$

$$P_2 = P_1 + V_2(q_2) - V_2(q_1)$$

$$P_3 = P_2 + V_3(q_3) - V_3(q_2)$$

In this case, in the second period, the firm implements $q_4 = \theta_4$, $q_3 = 2\theta_3 - \theta_4$, $q_2 = \frac{5\theta_2 - 2\theta_3}{3}$ and $q_1 = \frac{4\theta_1 - \theta_2}{3}$. The second period profits are:

$$\Pi^2_{123|4} \;=\; \frac{(\theta_4)^2 + (2\theta_3 - \theta_4)^2 + (\frac{5\theta_2 - 2\theta_3}{3})^2 + (\frac{4\theta_1 - \theta_2}{3})^2}{4}$$

The surplus of the lowest type is:

$$\underline{U}_{123|4} \;=\; V_1(q_1) - P_1 \;=\; V_1(q_1) - \frac{V_1(q_1) + V_2(q_1)}{3}$$

(I think the expression above is wrong!)

This can be simplified (this is a new expression):

$$\underline{U}_{123|4} \;=\; -\frac{2}{3}(V_2(q_1) - V_1(q_1)) - \frac{1}{3}(V_3(q_2) - V_2(q_2))$$

8.7 Both pairs $\{\theta_1, \theta_2\}$ and $\{\theta_3, \theta_4\}$ are uninformed case

That is, $\tilde{m} = \{\ \{\theta_1, \theta_2\}, \{\theta_3, \theta_4\}\ \}$.

Using similar steps as above, the quantities offered are: $q_4 = \theta_4$, $q_3 = \theta_3$, $q_2 = 3\theta_2 - \theta_3 - \theta_4$ and $q_1 = 2\theta_1 - \theta_2$ (if monotonicity constraint was not binding)

Unfortunately, MC is binding for the last types which means that $q_1 = q_2$. This means that the profits are given by:

$$\Pi^2_{12|34} = \frac{(\theta_4)^2 + (\theta_3)^2 + (3\theta_2 - \theta_3 - \theta_4)^2 + (2\theta_1 - \theta_2)^2 - \Delta^2}{4}$$

Where I used that $W(q, \theta) = \theta^2 - (q - \theta)^2$ to derive the cost of distortion from M.C.

9 Appendix 2 - first period profits

As defined above, $\Pi^1_{\tilde{m}}$ are the profits from a first period contract which generates (through passive memory) a memory partition $\tilde{m}$ in the second period. For example, when $\tilde{m} = \{ \{\theta_1, \theta_2, \theta_3\}, \{\theta_4\} \}$ these profits are going to be written as $\Pi^1_{123|4}$.

In the first period, the firm can extract all the consumer surplus. This means that the firm will maximise the total welfare[15] of the contract. The total welfare of a contract, for a consumer with type θ who is offered quantity q, can be written as[16]:[17]

$$W(\theta, q) \;=\; 2\theta q - q^2 \;=\; (q^*(\theta))^2 - (q - q^*(\theta))^2$$

Where $q^*(\theta)$ is the welfare maximising quantity such that $q^*(\theta) = \theta$ in this case.

In general, the allocation offered by the firm in the first period is going to differ from the welfare maximising by the fact that the firm might want to distort the contracts to prevent consumer learning. I will say that the firm *prevents learning of type* θ_1 *and* θ_2 when $m(\theta_1) = \{\theta_1, \theta_2\}$. And more generally I will say that the firm chooses a *pricing which prevents consumer learning* when the firm is *preventing learning of some types* as defined above.

Because of the passive memory, when preventing learning of any two consecutive types[18] θ_i and θ_{i+1} the firm needs to set $P(\theta_i) = P(\theta_{i+1})$. Because the consumers have utility that is always increasing in the quantity consumed q this means that the firm has also to choose $q(\theta_i) = q(\theta_{i+1})$.

You can use the allocation above to see that, in general, when preventing learning of consecutive types the firm is going to offer a pooling quantity that is equal to the expected type in the set of types that are being prevented to learn.

[15] Where welfare is defined as the sum of the firm's profits and the consumer's utility

[16] This is a consequence of the linear utility and quadratic cost function

[17] This fact does not depend on constants. For example, the expression would be the same $W(\theta, q) = (q^*(\theta))^2 - (q - q^*(\theta))^2$ if $W(\theta, q) \;=\; \theta q - q^2$. With the only difference being that $q^*(\theta) = \frac{\theta}{2}$

[18] Preventing learning of non-consecutive types - in the sense of having types with the same (partially uninformed) posterior - is not possible because the firm would need to have decreasing prices.

That is, when preventing learning of types in Θ^P such that $\Theta^P \subset \Theta$ the firm offers $\bar{q}_{\Theta^P} = E[\theta|\theta \in \Theta^P]$

When not preventing learning the firm will offer quantities that maximise the welfare, that is $q(\theta) = q^*(\theta) = \theta$

For example:

$$\Pi^1_{12|3|4} = \frac{2(\frac{\theta_1+\theta_2}{2})^2 + (\theta_3)^2 + (\theta_4)^2}{4}$$

10 Proofs:

10.1 Pooling profits always exceed screening profits

Here, I establish that profits when pooling all types are higher than when screening all types when the welfare function is quadratic (linear utility and quadratic costs) and the firm sells to all consumer types.

Lemma 8. *Pooling profits are higher than screening ones, such that:* $\Pi^1_{1234} > \Pi^2_{1|2|3|4}$

Pooling case The pooling profits are given by welfare. As stated above, with linear utility function and quadratic costs the welfare of a full pooling contract, where $\bar{q} = E(\theta)$, is:

$$\Pi^* - E\big[\ (\theta - E(\theta))^2\ \big]\ =\ \Pi^* - \sigma^2$$

which manipulating the variance formula becomes:

$$\Pi^* - \underbrace{\big[\ E(\theta^2) - \ \big(E(\theta)\ \big)^2\ \big]}_{\text{Quantity distortion}} \tag{1}$$

Screening case The screening profits are given by welfare minus information rents. In general, $q^S(\theta) = \theta - \frac{1-F(\theta)}{f(\theta)}$. Which means that there is *no shutdown* of any type when:

$$\theta > \frac{1 - F(\theta)}{f(\theta)}$$

for all $\theta \in \Theta$. Then, profits when screening are:

$$E\big[V - C - V_\theta \frac{1 - F(\theta)}{f(\theta)}\big]$$

or written as cost of distortion and information rents:

$$\Pi^* - \underbrace{E\big[\ (\theta - q^S(\theta))^2\ \big]}_{\text{Quantity distortion}} - \underbrace{E\big[\ 2q^S(\theta)\frac{1 - F(\theta)}{f(\theta)}\ \big]}_{\text{Info rents}}$$

Substitute $q^S(\theta)$ to get this formula:

$$\Pi^* - \big(\; E\big[2\theta \frac{1-F(\theta)}{f(\theta)}\big] - E\big[\; (\frac{1-F(\theta)}{f(\theta)})^2\big] \; \big) \tag{2}$$

Comparison Equation 1 exceeds equation 2 whenever:

$$E[\; \big(\theta - \frac{1-F(\theta)}{f(\theta)}\big)^2 \;] > [\; E(\theta) \;]^2 \tag{3}$$

The condition given by equation 3 is always true for uniform distribution. Note that Jensen's inequality is what allows this to be, possibly, positive.

IMPLICATION 1 This implies, in this draft's notation, that pooling contract dominates the myopic contract since:

$$\Pi^1_{1234} + \Pi^2_{1234} \; > \; \Pi^1_{1|2|3|4} + \Pi^2_{1|2|3|4}$$

given that $\Pi^1_{1|2|3|4} = \Pi^2_{1234}$

10.2 Intuition for previous result and an example with two types

- Costs of pooling are just in terms of quantity distortions since the consumers are not informed. In this quadratic framework, these costs are just the squared distance to each type optimal quantity θ.

  In two type case, this is $(\frac{\Delta}{2})^2$.

- Costs of screening are the distortion in quantities plus the information rents that the firm has to pay to prevent deviation. A first simple step is to see that, if the firm did not distort quantities, the cost would be only the (non-optimal) information rents.

  With two types these are $V_2(q_1) - V_1(q_1)$. But given that optimal q_1 is θ_1 this becomes $2\theta_1(\theta_2 - \theta_1)$. Or using equidistant types $2\theta_1\Delta$

- In the example, the costs of screening using optimal quantity ($q_1^S = 2\theta_1 - \theta_2 \; = \; \theta - \Delta$) are distortion costs $\frac{\Delta^2}{2}$ and information rents $2(\theta - \Delta)(\theta_2 - \theta_1) = (\theta - \Delta)(\Delta)$ (both costs are multiplied by probability 1/2).

- For screening to be profitable the lowest type has to be higher than width of interval of types Δ (for more than two types this condition is stronger).[19]

  In the example with two types, $\theta_1 > \Delta$.

- This means that costs of pooling are smaller than costs of screening.

10.3 Pooling contract has higher profits than any other contract

1) Previous subsection proves that pooling contract dominates the myopic contract.

2) Previous section also proves that when pooling for all types pooling gives higher profits than screening. This is also going to be true for comparisons over smaller intervals of types. The proof follows from the one above.

3) Only issue I guess is if smaller intervals with pooling have profits that are high enough to compensate learning in the second period.

10.4 Surpluses (of lowest type) are increasing in learning

10.5 3PTs have higher profits than inverted 3PTs

[19] Note that if screening is not profitable it might also be optimal to drop some types of pooling contracts.