Information Transmission and Recommender Systems

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Deran Ozmen

Dissertation Director: Dirk Bergemann

December 2005

UMI Number: 3194693

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

Information Transmission and Recommender Systems

Deran Ozmen

2005

The first chapter of this dissertation studies optimal pricing in the presence of recommender systems. A recommender system is a program that internet sellers employ to infer what product a returning customer would like using data on the feedback from all customers' past purchases. Its effect on the market is twofold: (i) it creates value by reducing product uncertainty for the customers and hence (ii) its recommendations can be offered as add-ons, which generates informational externalities as the quality of the recommendation add-on is endogenously determined by sales. The chapter investigates the impact of these factors on optimal pricing of different products for a seller with a recommender system in the presence of a competitive fringe without such a system. The main finding is that, if the recommender system is sufficiently effective in reducing uncertainty, the seller prices otherwise symmetric products differently to have some products experienced more aggressively.

The second chapter analyzes an information transmission problem between multiple parties. An exogenously informed party tries to convey his information to two uninformed audiences, who react to the same news in opposite ways. In this setting, satisfying one audience might come at the expense of upsetting the other. The chapter focuses on this dilemma to provide insights into why information is shared privately behind closed doors in some cases and is made public in others.

The third chapter studies the problem of a monopolist selling a durable good over time. The model deviates from the standard dynamic monopoly models by assuming that the monopolist can vary the quality of the product each period and that the buyers are differentiated with respect to their valuation of quality. In this setting, the chapter investigates to what degree the Coase conjecture applies, given that the buyers who are willing to wait for low prices know that low prices will possibly come with low quality. The finding is that despite of this low quality threat, the seller still gets zero profits at the limit.

Acknowledgments

I would especially like to thank Dirk Bergemann for invaluable discussions, feedback and encouragement. I would also to thank Benjamin Polak, Stephen Morris, Dino Gerardi, Judy Chevalier, Dina Mayzlin and seminar participants of the Graduate Summer Workshop and the Game Theory Lunch for many helpful comments and conversations.

To my parents Ruya and Bedrettin,

in whom I've found endless support and love...

Contents

List of Tables

List of Figures

Chapter 1

Optimal Pricing Policy with Recommender Systems

1.1 Introduction

New developments in computer technology and the increased usage of internet by customers create new questions to be analyzed by economists. Given the large volume of transactions on the internet, it is natural to analyze its differences from brick-and-mortar markets. This paper's central interest is one such difference, the fact that internet sales pave way to a large accumulation of data about customers and products. Internet sellers can easily build large databases that consist of personalized data on all their customers, the customers' past purchases and the feedback from those purchases. In this paper we analyze one particular use for the information accumulated in these databases, "recommender systems". A recommender system is a software program which uses the accumulated data to make statistical inferences about what product a particular customer would like when she returns to the website. The best example of such a system is that employed by Amazon.com.

Once a customer makes a purchase there, the next time she logs on to Amazon.com, a recommendation pops up on the screen for her. There are many other internet sellers, such as CDNOW.com, Reel.com, Netflix.com, MovieLens.org, that employ some version of a recommender system.

From an economic point of view, a recommender system represents an informational linkage that creates additional surplus by reducing uncertainty for the customers. In this paper we present a two-period, two-product model that describes the interaction between a seller employing a simple recommender system and a competitive fringe with no such system, to analyze the surplus created by recommender system and the different dynamics it generates in the market.

There are usually two sources of uncertainty involved in the decision process of a customer. She may be unsure about her tastes and/or characteristics of the products. In our model, we focus only on product uncertainty in the on-line market for horizontally differentiated products, where the difference in customers' tastes translate into differences in the willingness to pay for decreased uncertainty. Our recommender system acts as a mechanism that collects customer evaluations, through which the seller infers more information about the products. Rather than modelling the evaluation process for each customer, we employ an information structure that aggregates these evaluations into a single signal that the seller receives on each product. The seller reveals whatever inference he makes to his "loyal" customers, those who have made a purchase from him before. Thus, a loyal customer has the chance to make a better informed choice using the inference revealed to her by the recommender system.

The surplus created by the recommender system can be directed to increase sales and/or increase prices. The possibility of increasing sales has been documented by Chevalier and Mayzlin (2003). They empirically investigate the impact of customer reviews on sales of books in Amazon.com and BarnesandNoble.com. They find that the relative market share of a book across the two sites is related to differences across the sites in the number of reviews for the book. This enforces the idea that the volume of reviews has a positive impact on sales. The possibility of an extraction through prices arises due to the loyalty factor mentioned above. Future recommendations might be considered as add-ons to current purchases from a seller with recommender system. Hence buyers may agree to pay higher prices for the products they purchase from a seller with a recommender system today so that they can receive recommendations in the future. Brynjolfsson and Smith (2001)'s empirical investigation of consumer behavior at internet shopbots for books provide evidence for existence of such behavior by consumers. They find that online book buyers are willing to pay a positive premium to purchase from the sellers they have either visited or shopped at before. One interpretation of this premium is that it is the fee for the information the sellers sell through the recommender system to loyal customers. These empirical facts can support the role of recommender system in increasing both the sales and prices. In this paper we focus on the latter by assuming that each buyer has unit demand each period. Hence optimal pricing is the main focus of the paper. We seek to answer how much of the surplus a seller with a recommender system can extract from customers through pricing in the presence of a competitive fringe.

As we mentioned above, a recommendation can be considered as an add-on: it is an additional service a customer receives on top of the purchase she makes. Another interpretation on the same lines is that, future recommendations are information goods that are bundled with current purchases. The recommendations and products form pure bundles as defined by Adams and Yellen (1976): it is not possible to purchase the bundle elements separately. Recommendations, however, are different from typical add-ons and bundle elements because their quality is determined endogenously by the information accumulated through the seller's sales. Thus the seller's pricing problem incorporates the additional need to set the quality of the add-on for each product optimally, which is equivalent to gathering the optimal amount of information on each product. Therefore the seller's dual problem of what market share to capture and how to distribute the buyers over different products entails informational externalities. These externalities can be separated into two elements. The first element is what we call the "volume externality". This externality represents the general coordination element inherent in the problem, which is that as the seller has more customers, he will be able to make better recommendations and thus attract more customers. This element determines how much of the market the seller would like to capture. The second one is the "product externality". This externality relates to the distribution of buyers within one seller over different products. If there are a lot of customers buying one particular product in one period, others may be willing to delay the purchase of that product and be directed to other products for that period. The strength of this effect determines whether the seller tries to accumulate equal amounts of information on each product or whether there are increasing returns to information so that the seller tries to induce large volume of buyers to buy some products and provide information at the expense of other products on which smaller volume of information is gathered.

The volume and product externalities become stronger as the recommender system performs better in reducing uncertainty. Not surprisingly, the recommender system's performance increases in the degree of uncertainty about the products and in the precision of customers' evaluations. More interestingly an increase in this performance leads to changes in the seller's pricing policy and segmentation of the market.

We find that when the recommender system does not perform well in reducing uncertainty, the seller prefers to gather equal amount of information on symmetric products by pricing them uniformly. The buyers with sufficiently high willingness to pay for reduced uncertainty agree to pay this price to benefit from the recommendation service and the others simply decline this service and purchase from the fringe. For example consider the books market. Suppose two novels "Double Homicide" and "The Rocky Road to Romance"[1] are introduced for sale at the same time. It is very clear that the first one is a mystery and the second one is a romance novel. Hence there will not be many buyers willing to pay a premium to receive information on the type of either novel. There is not much the seller can gain by speeding up the information accumulation, hence he prices the products similarly.

Our results show that as the performance of the recommender system increases, the seller implements differential pricing which segments the market such that some products are experienced by a larger group of buyers than others. The buyers with high willingness to pay for reduced uncertainty choose to be in the smaller group to be able to use the information provided by the large group. Those with the low willingness to pay choose to

[1]These novels are new releases that can be found on Amazon.com.

be in the large group to benefit from the lower price. In some cases, this price is so low that it implies a loss for the seller on that particular product. The seller is willing to bear this loss because the information gathered allows him to subsidize it through sufficiently higher prices on other products. Let us consider the books market again. Suppose "The Syme Papers" and "Jonathan Strange & Mr. Norrell: A Novel"[2] are both new releases by new authors. These titles clearly do not reveal any relevant information about the type of these books. Customers who are very particular about the type of book they read would be willing to pay a premium for more information before they make their purchases. To extract this premium the seller needs to gather enough information on at least one book. Hence he targets one of the two books, charges a lower price for that to speed up the information accumulation.

We investigate the segmentation in the market further. In our model, the customers differ both in the type of product they prefer and also in the intensity of their preference. Some buyers are more flexible in their choices than others. It is the buyers with inflexible tastes who really benefit from the recommendation service. The interesting question then becomes whether the seller segments customers of one type of product from the customers of the other type or whether he segments the inflexible customers of both types from the flexible buyers. We find that the former kind of segmentation occurs when the recommender system has a low performance and the latter occurs when it has a high performance.

We also ask whether or not the resulting allocation of buyers to different products and sellers as a result of the optimal pricing scheme is efficient. We find that the optimal pricing

[2]These books can be found on Amazon.com's website as well.

mechanism does not necessarily employ the recommender system at the efficient level. In particular, sometimes information is wasted, because the seller does not capture the whole market. We also find that the seller might over-utilize the system for some products and under-utilize for others leading to a more unequal information level obtained on different products then it should be.

The road map is as follows: We discuss the related literature in Section 1.2. Then we describe the model in Section 1.3. In Sections 1.4 and 1.5 we look at the efficient solution and the equilibrium respectively. Then in Section 1.6 we compare the efficient and equilibrium solutions. And finally we conclude by discussing the contributions of the recommender system to society and the areas for future research.

1.2 Related Literature

In the literature, the only formal model related to recommender systems is introduced by Avery, Resnick and Zeckhauser (1999). They take a mechanism design point of view towards the problem and focus on designing a pricing/subsidy mechanism to induce efficient provision of evaluations. Avery and Zeckhauser (1997) gives a less formal description of the same problem through some examples. In their model, there is a single product which can be of two types, "Good" or "Bad". There is a set of agents trying to decide whether to consume the product or not. If the product is good they get a positive utility and if the product is bad they get a negative utility. The agents are differentiated with respect to the utilities they derive from the two types of the product. The agents who consume the product earlier provide (honest) imperfect evaluations on the type of the product to the agents who have not consumed it yet. In this setting, the efficient provision of evaluations refers to both the optimal quantity of evaluations and also the optimal sequencing of evaluations given the differences in the agents' utilities. They show that the efficient provision can be achieved through a broker who offers the agents side payments to choose the socially optimal actions. In particular, the payment schedule involves subsidies to the agents who produce evaluations earlier in the game. However, there is no consideration of profits on the part of the intermediate agent who implements the pricing mechanism to generate the efficient allocation. In this paper, we take the mechanism as given and try to maximize profits. The sequencing of agents over time in their model resembles the idea of distributing buyers over different products in this paper. In both cases, the buyers who provide more information are somehow subsidized.

Shapiro and Varian (1999) and Vulkan (2003) both analyze the novelties induced by the internet and the sharing of information in e-commerce. Varian and Resnick (1997) give a brief description of recommender systems and the issues they raise. In particular they are concerned with the privacy and incentive problems that are brought about by the recommender systems. They also explain that the larger the customer base of a recommender system, the more customers would be willing to use it, which is equivalent to what we earlier described as the "volume externality".

Chevalier and Mayzlin, as mentioned before, empirically analyze the relation between customer reviews and sales in Amazon.com and BarnesandNoble.com. They use sales and customer review data publicly available on the two websites. They characterize the rating behavior of the customers and they find that the reviews are predominantly positive at both websites. They estimate the effect of both the number of reviews and the content of the reviews on the relative sales of each website. They find that the number of reviews on a site have a positive impact on the sales of that site and moreover an improvement in a book's review leads to an increase in relative sales at that site.

Brynjolfsson and Smith (2001) use data from EvenBetter.com, an internet shopbot selling books, to analyze the shopping behavior of consumers. In particular they employ multinomial logit and nested logit models to estimate how the consumers respond to brand name, retailer loyalty, prices, and contractible and non-contractible product characteristics. They find that the three brand names Amazon.com, BarnesandNoble.com and Borders.com on average have a \$1.13 price advantage for books that sell for \$36.80 on average. As we mentioned earlier, they also find evidence supporting consumer loyalty. They specifically calculate that retailers that a consumer had selected previously on the shopbot hold a \$2.49 advantage over other retailers.

As we mentioned before, the analysis of recommender systems inherits some features from the literature on product add-ons and multi-product bundling. In the literature there have been many different reasons given to why a monopolist might prefer to bundle his products. Eppen, Hanson and Martin (1991), Adams and Yellen (1976), Schmalensee (1984) suggest reasons such as cost savings, complementarities between different products or extraction of more consumer surplus as there will be less diversity in the valuations of the consumers for the bundles compared to the valuations for individual products. In our model there is a strategic reason behind bundling. The seller is offering an information good not provided by his competitors. However, the value of this good depends on the volume of his sales of the main product. If he unbundles, on the product side he might lose buyers to other sellers which decreases the value of the information good he is offering. Our results and methodology would apply to more general settings that involve competitive sales of pure bundles, where the value of at least one element in the bundle is determined by the overall sales.

The growing seller interest in recommender systems led to many computer scientists and IT researchers working on this topic with a technical focus on writing the most efficient recommender system. In computer science, the recommender systems we discuss here are formally known as "collaborative filtering systems". The system keeps a database in the form of a matrix which has the customer's identity in the rows, the products in the columns and the ratings received from customers as the elements. The collaborative recommender system predicts ratings for the products that have not been purchased by a customer based on the products previously rated by other customers. The system first computes a similarity (correlation) measure between customers. The rating estimate for a particular customer-product couple is the average rating left by other customers weighted by the similarity measures. There are other collaborative filtering methods as well. For example as published in the "IEEE Internet Computing: Industry Report", Amazon.com uses a modified collaborative filtering method, which they refer to as the "item-to-item based collaborative filtering". Their method computes similarity measure between the items rather than the customers and then recommends the items similar to what a customer has purchased before. Breese, Heckerman and Kadie (1998), Mild and Natter (2001), and Ansari, Essegaier and Kohli (2000) describe and compare other methods of prediction which range from Bayesian methods of estimation to regressional methods. In all these cases the physical procedure of making use of other customers ratings to make a recommendation to a customer explicitly reveals how the externality is incorporated into the problem. In this paper we take the collaborative recommender system as given and model the recommender system so as to generate some of the externality effects inherent in collaborative filtering.

1.3 The Model

In this section we introduce a two-period model where a seller with a recommender system and a fringe with no such system compete in prices in a market for horizontally differentiated products. In this market there are two types of the product and a continuum of buyers. In period 0 two different products are offered by the sellers. The sellers and the buyers share a common prior about the type of each product. These products are differentiated only with respect to the prior they arrive with. Each buyer chooses a product to buy and a seller to buy from in period 0. The seller with recommender system collects information from his customers about the products purchased from him in period 0. In the second period, he reveals this information as recommendations to the buyers who purchased from him in period 0. In the second period a new product arrives at all sellers and buyers again choose a product and a seller to buy from given their recommendations. The following figure gives the time-line before we go into the details of the model.

1.3.1 The Market

There is one seller with a recommender system, denoted by M, and a competitive fringe with no such system, denoted by F, in the on-line market for a particular product group.

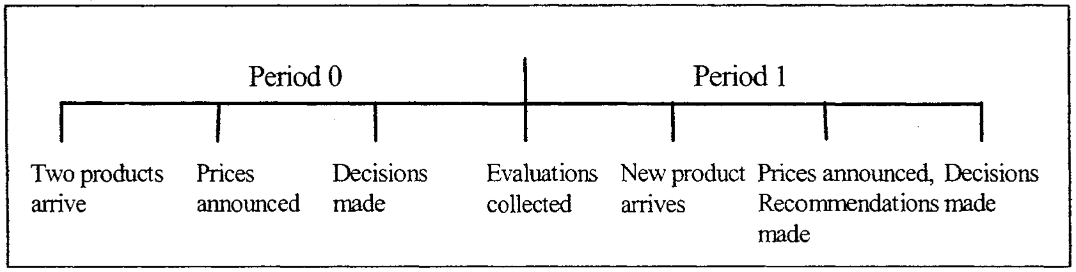

Figure 1.1: The timeline

Within the market, there are two different types of the product, denoted by $x \in \{-1, 1\}$. There is a continuum of buyers in $[-1, 1]$ with unit mass, where each buyer is characterized by his preference $\theta \in [-1, 1]$. θ is distributed uniformly in $[-1, 1]$. The gross utility a buyer of type θ derives from a type x product is specified as

$$u(\theta, x) = v - (\theta - x)^2 \tag{1.1}$$

As an example consider the product line to be books. Then the two types of the product can represent "mystery" versus "romance" novels. We can consider the buyers with preference parameters close to -1 or 1 as "inflexible" and buyers with preference parameter close to 0 as "flexible", because the former group would insist on their favorite kind of book whereas the latter group would not be adverse to trying other kinds. In a more general context, it is the former group who has more to lose if they get a product with a type further from their taste, whereas the latter group's utility decreases by little in that situation. This means the flexible buyers could potentially be the experimenters of new products if they are given enough incentives. We will see in the following sections that the seller will exploit this feature.

1.3.2 Timing and Choices

There are two periods with flow of products and there is uncertainty about their types. The sellers and buyers share a common prior on these products' types. In period 0 two products arrive at all sellers denoted by l and h. These products are differentiated only with respect to the priors attached to them. Let $x_i \in \{-1, 1\}$ be the true type of product $i \in \{l, h\}$ and $\alpha_i \equiv \Pr(x_i = 1)$. We assume that the two products arrive with symmetric uncertainty, i.e.

$$\begin{aligned} \alpha_h &= \frac{1}{2} + \varepsilon \\ \alpha_l &= \frac{1}{2} - \varepsilon \end{aligned} \tag{1.2}$$

where $\varepsilon \in \left[0, \frac{1}{2}\right]$. Hence the initial priors are differentiated by ε, which we will refer to as the "initial information".

In period 1 a new product, m, arrives with prior $\alpha_m \in \{\alpha_l, \alpha_h\}$ at all sellers. In period 0, neither the buyers nor the sellers know the exact value of α_m, but they attach $\frac{1}{2}$ probability to α_m being α_h and α_l.

1.3.3 Payoffs

Marginal cost of each product for all sellers is c. We assume that the price for each product in the competitive fringe equals c. There is no discounting. Each buyer buys at most one product each period. Moreover, a buyer has to buy a different product each period. We also assume that per period outside utility for each buyer is smaller than $V - c - 4$, so that each buyer is willing to buy some product each period from the fringe. Hence, in our model fringe is the buyers' outside option.

1.3.4 Learning Through the Recommender System

Between periods 0 and 1 seller M receives information form his buyers. We aggregate the information as follows: Let μ_i denote the measure of buyers who buy product $i \in \{l, h\}$ from seller M in period 0. Seller M receives a random signal $y_i(x_i) \in \{\emptyset, -1, 1\}$ on the type of each product $i \in \{l, h\}$ between periods 0 and 1, where

$$\begin{aligned} \Pr(y_i(x_i) = \emptyset \mid x_i) &= 1 - \mu_i \\ \Pr(y_i(x_i) \in \{-1, 1\} \mid x_i) &= \mu_i \end{aligned}$$

We can interpret a signal of $\emptyset$ as no signal. Given that the seller receives a relevant signal, the probability of the signal being correct is

$$\Pr(y_i(x_i) = x_i \mid y_i(x_i) \in \{-1, 1\}, x_i) = \frac{1}{2} + \gamma$$

where $\gamma \in \left[0, \frac{1}{2}\right]$. Therefore γ can be interpreted as the informativeness of the signal when received. The event tree in Figure 1.2 summarizes the signal structure where $x_i' \neq x_i$.

Given this random structure, we assume that the recommender system is a pre-committed direct mechanism that computes the posterior beliefs for each product i based on the signal y_i and reports them only to the buyers who have bought from him in period 0. The posterior for product i given signal y_i will be denoted by $\alpha_i'(y_i) \equiv \Pr(x_i = 1 \mid y_i)$.

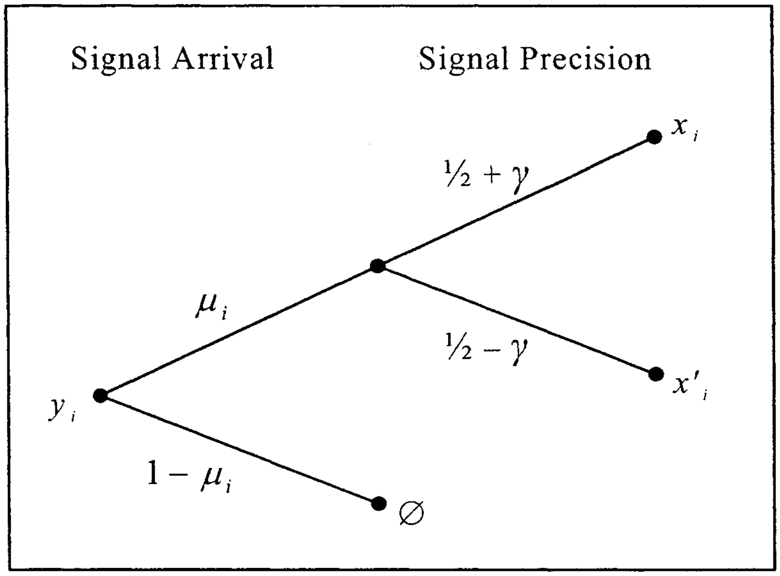

Figure 1.2: The Signal Structure

1.3.5 Pricing

In period 0, seller M announces prices for each product, i.e. $\mathbf{p} = (p_l, p_h) \in \mathbf{R}^2$. The search cost is zero for all buyers, thus each buyer logs onto all websites and observes all prices, and then simultaneously chooses a product to purchase $i \in \{l, h\}$ and a seller to buy from $s \in \{M, F\}$.

In period 1, seller M announces prices for each product, $(p_l', p_m', p_h') \in \mathbf{R}^3$ and reveals the recommendations to the buyers who have purchased from him in period 0. Let j denote $\{l, h\} / i$. Then buyers simultaneously choose a product to purchase, $i' \in \{j, m\}$, and a seller to buy from, $s' \in \{M, F\}$, given the information they observe. Notice that a buyer can get the recommendation from seller M and still purchase from the fringe in period 1, because search costs are zero.

1.3.6 Interpretation of the Model

There are two products arriving with symmetric uncertainty attached in period 0. A high ε means there is less uncertainty about each product's type and that the two products are highly differentiated. A low ε means uncertainty is high for both products and that initially the two products look similar. In terms of the books example, a high ε would mean that either the books have very revealing titles or the authors' styles are very well known. Similarly a low ε can be generated by very vague titles and/or new authors.

Through the signal structure we described in Figure 1.2, seller M gains information about the type of the two products. Suppose a buyer buys product i from seller M in period 0. Then in period 1 her choice set is $\{j, m\}$. The recommender system supplies information to the buyer about j's type. Hence the buyer can make a better informed choice between j and m. This describes the contribution of recommender system and how it creates additional surplus. The extent of this contribution depends on ε and γ. Let

$$\rho = \frac{\gamma}{\varepsilon} \tag{1.3}$$

and we interpret ρ as the "performance of the recommender system". The reason is that when γ is high the signals are more precise and thus the updating will be more critical for the buyers' choices, and when ε is low, the products are too unknown and any new information is very valuable. Thus a high ρ actually increases the effect of recommender system in reducing uncertainty. One interpretation of γ is that it is the likelihood that the buyers will leave true evaluations. Hence it will be higher for markets in which a higher proportion of buyers tend to leave evaluations and there are hardly any incentives to lie. An article in New York Times on February 14, 2004 entitled "Amazon Glitch Unmasks War Of Reviewers " drew attention to the fact that in the books and CDs market a non-negligible share of the evaluations might be generated by the authors and the singers themselves and their friends or their enemies. This kind of knowledge decreases the customers' beliefs in the recommender system and hence lowers γ.

Figure 1.2 shows that the probability of receiving a signal on a product increases in the measure of buyers buying that product. This captures the effect that as a seller has more customers, the recommender system will have more input and make better recommendations.[3] In the next section we will show that it is this event tree combined with the preferences that generates the "volume externality effect" and "product externality effect" that we mentioned in the introduction.

1.4 Efficiency with the recommender system

In this section we analyze the basic problem that we face in terms of efficiency, i.e., how to distribute buyers over different sellers and products to maximize total surplus. We first introduce new definitions that summarize the important issues regarding efficiency. Then we display the formal characterizations and discuss the trade-offs involved in the maximization of total surplus. Finally we characterize the efficient solution and also look into comparative statics.

[3] An alternative way of modelling this effect would be to assume that the sender receives a signal on each product with probability 1, but the precision of the signal depends on the measure of buyers buying that product, i.e.we can the consider the precision of the signal received on a version as $\gamma_i(\mu_i)$. This functional form has the property that $\gamma_i(\mu_i)$ increases in μ_i and $\gamma_i(0) = 0$ and $\gamma_i(1) = 1$. The alternative does not change the qualitative results in any significant manner. For quantitative simplicity, we therefore have assumed the structure described in the section, which provides us linearity in γ_i.

First notice that, period 1 does not pose an interesting question regarding efficiency. There are no future considerations, so efficiency requires each buyer to purchase the product that maximizes her expected utility given the information she receives from seller M's recommender system. Period 0 poses a more challenging puzzle.

The efficient allocation in period 0 in the absence of a recommender system is straight forward. If there is no recommender system, or equivalently if $\varepsilon = \frac{1}{2}$ or $\gamma = 0$, there are no informational returns, because there is either no product uncertainty or the seller's signal is uninformative. Thus, each buyer should be allocated to the product that gives her the highest per period utility, i.e. all buyers with $\theta \geqslant 0$ should buy product h and all buyers with $\theta < 0$ should buy version l. Notice that the seller choice does not matter in this case, because there is no difference between the service provided by different sellers.

If we introduce some uncertainty and informativeness into the setting, it is no longer true that each buyer should buy the product which gives her the highest per period expected utility, because a buyer's choice of seller and product in period 0 affects the utility of all the other buyers in period 1. In particular, buyers' purchases in period 0 generate a trade-off between two effects: A direct effect on their utility in period 0 and an indirect effect on the utility of all buyers in period 1 due to the informational externality generated by the recommender system, which is evident in Figure 1.2. We can split this externality into volume externality and product externality as we did in the introduction. There are two variables of importance, which incorporate these externalities respectively : (i) the distribution of buyers over sellers M and F and (ii) the distribution of buyers over the two different products. The first variable is important because it determines the aggregate information gathered by seller M and thus the overall effectiveness of the recommender system in reducing uncertainty. It is clear that all buyers should purchase from seller M in period 0 because information has positive value and the inflow of information is maximized when seller M has full market share. Therefore, we ignore the fringe in our remaining analysis of efficiency. The second variable is important because it is a potential way to create endogenous differentiation between the two products which are exogenously differentiated by ε. In particular, if the measure of buyers purchasing each product from seller M is different, the products will be differentiated even further when the posteriors are computed. We introduce the following definition to refer to this endogenous differentiation.

Definition 1.1 (BALANCE). *A distribution of buyers with (μ_h, μ_l) is balanced if $\mu_h = \mu_l$ and unbalanced if $\mu_i > \mu_j$ for some $i \in \{l, h\}$. The degree of unbalance is given by $\frac{\mu_i}{\mu_j}$.*

We are concerned with whether it is efficient to create endogenous differentiation through an unbalanced distribution and if so, which buyers should benefit from such an unbalance. In other words, if the distribution is unbalanced, one product is experimented by a larger group of buyers and the small group of buyers wait to benefit from their feedback. If this is the case, then it is also important for efficiency to know the composition of these groups. Regarding this last point we introduce the following definition.

Definition 1.2 (SORTING). *For a given distribution of buyers with (μ_l, μ_h), we say the distribution is*

1. *"sorted" if the set of buyers buying products l and h respectively are line segments of the form $[-1, \cdot]$ and $[\cdot, 1]$,*

2. *"shuffled" if* $\mu_i \geqslant \mu_j$ *for some* $i \in \{l, h\}$ *and the set of buyers buying product* j *consists of two segments* S^-, S^+ *of the forms* $[-1, \cdot]$, $[\cdot, 1]$ *respectively, where* $\frac{\min\{|S^-|,|S^+|\}}{\max\{|S^-|,|S^+|\}}$ *is the degree of shuffling and*

3. *"perfectly shuffled" if* $|S^-| = |S^+|$.

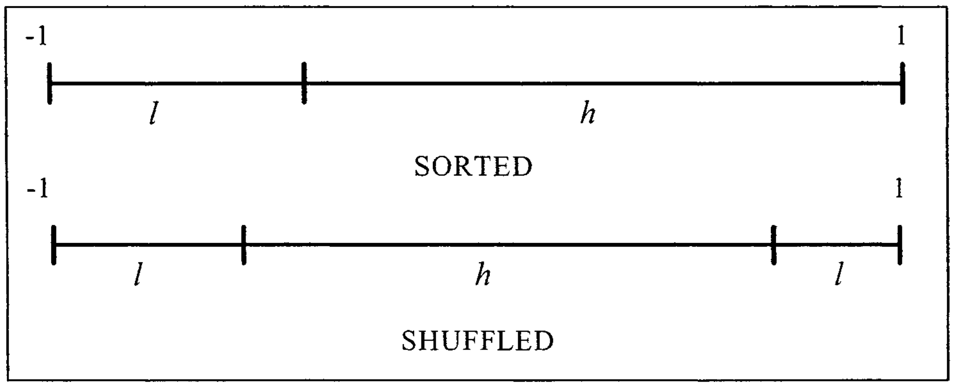

Figure 1.3: Sorted and mixed distributions for $\mu_h > \mu_l$

Figure 1.3 illustrates Definition 1.2. If a distribution is shuffled, it is the inflexible buyers of both types that benefit more from the endogenous differentiation created by the unbalanced distribution. In other words, one product is experimented by a large group of flexible buyers and the inflexible buyers of both types receive good recommendations from the experiences of the former group. On the other hand, if a distribution is sorted, it is usually the inflexible buyers of one type receiving information from the experiences of all other buyers.

With these definitions in mind, we first characterize the per-period expected utility and then the two-period value function for each buyer. The per period expected gross utility for a buyer of type θ from purchasing product $i \in \{l, m, h\}$ given α_i is

$$\begin{aligned} \mathbb{E}_{\alpha_i} u(\theta, x_i) &= v - \alpha_i (\theta - 1)^2 - (1 - \alpha_i)(\theta + 1)^2 \\ &= v - (\theta + 1)^2 + 4\alpha_i \theta \end{aligned} \tag{1.4}$$

First, as we discussed above, the efficient period 1 allocation maximizes the period 1 expected utility with respect to the remaining choice set $\{j, m\}$, given the recommendations. Second, notice that equation (1.4) is linear in α_i. These two facts imply that, from a period 0 point of view, it is the *expected maximal (minimal) posterior* that determines the expected utility of a buyer of type $\theta \geqslant 0$ ($\theta < 0$) in period 1. These posteriors for a buyer who purchases product i in period 0 can be defined as

$$\begin{aligned} \overline{\alpha}_i(\mu_j) &\equiv \mathbb{E}\left(\max\left\{\alpha_m, \alpha_j'(y_j)\right\} \mid \mu_j\right) \\ \underline{\alpha}_i(\mu_j) &\equiv \mathbb{E}\left(\min\left\{\alpha_m, \alpha_j'(y_j)\right\} \mid \mu_j\right) \end{aligned} \tag{1.5}$$

A buyer with type $\theta \geqslant 0$ who buys product i in period 0 expects to get a product with this posterior in period 1 knowing that she will choose the product with highest probability of being type 1 once she receives information on j. A buyer with type $\theta \leqslant 0$ expects to get a product with a similar posterior, which this time is computed based on the fact that the buyer will choose the product with the lowest probability of being type 1 once she receives information.

Hence, the two-period gross value function for a buyer of type θ conditional on purchasing product i from seller M in period 0 is

$$U_M\left(\theta,i,\mu_j\right)=2v-2\left(\theta+1\right)^2+\begin{cases}4\theta\left(\alpha_i+\overline{\alpha}_i\left(\mu_j\right)\right) & \text{if } \theta\geqslant 0\\ 4\theta\left(\alpha_i+\underline{\alpha}_i\left(\mu_j\right)\right) & \text{if } \theta<0\end{cases} \tag{1.6}$$

The expected maximal posteriors conditional on purchasing each product can be derived as

$$\begin{aligned}\overline{\alpha}_h\left(\mu_l\right) &= \frac{1}{2}+\frac{1}{2}\mu_l\beta\left(\gamma,\varepsilon\right) \\ \overline{\alpha}_l\left(\mu_h\right) &= \frac{1}{2}+\varepsilon+\frac{1}{2}\mu_h\beta\left(\gamma,\varepsilon\right)\end{aligned} \tag{1.7}$$

where

$$\beta\left(\gamma,\varepsilon\right)=\begin{cases}2\gamma\left(\frac{1}{4}-\varepsilon^2\right) & \text{if } \gamma\leqslant\frac{2\varepsilon}{4\varepsilon^2+1}\\ \gamma-\varepsilon & \text{otherwise}\end{cases} \tag{1.8}$$

and the expected minimal posteriors can be derived symmetrically through $\underline{\alpha}_i\left(\mu_j\right)=1-\overline{\alpha}_j\left(\mu_j\right)$.

Equation (1.6) shows that the choice of a buyer in period 0 affects her expected utility in period 1 through the expected maximal (minimal) posterior given in Equation (1.7). Suppose a buyer of type $\theta>0$ purchases product h in period 0. She immediately receives the direct utility effect, α_h, in period 0. In period 1 her choice set will consist of $\{l,m\}$. She knows that she will receive information on product l with probability μ_l given the signal structure we described. Thus, in the event that she gets information, she will use it to choose optimally between l and m. The $\overline{\alpha}_h\left(\mu_l\right)$ term represents her expected maximizing choice which hence depends on μ_l. Simple computations show that the exact form of $\overline{\alpha}_h\left(\mu_l\right)$

is as given in equation (1.7). To gain some insights into this expression, suppose that the buyer decides to purchase m in period 1 regardless of the information. Since α_m is equally likely to be $\frac{1}{2}-\varepsilon$ and $\frac{1}{2}+\varepsilon$, in expectation she can guarantee a posterior of $\frac{1}{2}$, which is the first term in equation (1.7). But she knows that she will do better than that, because whenever the information she receives on l generates a posterior $\alpha_l' > \alpha_m$, she will choose l. Hence the second term represents his expected gain due to the information she will receive on l with a probability of μ_l. Notice that this expected gain increases in μ_l as it is more likely that the buyer will receive new information and thus make a better informed choice. Similarly if we look at $\overline{\alpha}_l(\mu_h)$, we can argue that the buyer can guarantee herself an expected posterior of $\frac{1}{2}+\varepsilon$ if she decides to purchase h regardless of the information she receives. But again, she can do better by purchasing m whenever she receives information such that $\alpha_h' < \alpha_m$. Hence the last term describes her expected gain due to the information she will receive with probability μ_h. The $\beta(\gamma,\varepsilon)$ parameter represents the informational gain and hence it depends on the likelihood that the information will be relevant, i.e. the likelihood that $\alpha_l' > \alpha_m$ and $\alpha_h' < \alpha_m$. For $\gamma \leqslant \frac{2\varepsilon}{4\varepsilon^2+1}$, the posteriors reveal the property that $\alpha_l'(1) \leqslant \alpha_h$ and $\alpha_h'(-1) \geqslant \alpha_l$, whereas this pattern is reversed for $\gamma > \frac{2\varepsilon}{4\varepsilon^2+1}$.

Notice from equation (1.6) that the two-period utility of a buyer with type $\theta > 0$ ($\theta < 0$) increases (decreases) with the expected maximal (minimal) posterior. Hence a buyer's preference over the two products may change with these posteriors as well. To allocate the buyers efficiently, we need to know the preference of each buyer, which clearly depends on both the initial prior and the expected maximal (minimal) posterior. The following lemma derives properties from equations (1.6) and (1.7), which describe how the preferences and in particular the expected posteriors are affected by informational changes. The properties are stated for buyers with positive types and expected maximal posterior, but the symmetric properties hold for buyers with negative types.

Lemma 1.1 (VALUE AND INFORMATION).

For all $i, j \in \{l, h\}$, $j \neq i$,

1. $\frac{\partial \overline{\alpha}_i(\mu_j)}{\partial \mu_j} = \frac{\partial \overline{\alpha}_j(\mu_i)}{\partial \mu_i} > 0$;

2. $\frac{\partial^2 \overline{\alpha}_i(\mu_j)}{\partial \mu_j \partial \gamma} \geqslant 0$ *and* $\frac{\partial^2 \overline{\alpha}_i(\mu_j)}{\partial \mu_j \partial \varepsilon} \leqslant 0$;

3. $U_M(\theta, i, \mu_j)$ *is supermodular with respect to* θ *and* $\overline{\alpha}_i + \alpha_i$;

4. *for all* $i', j' \in \{l, h\}$, $j' \neq i'$ *and* $\theta \geqslant 0$, $U_M(\theta, i, \mu_j) \geqslant U_M(\theta, i', \mu_{j'})$ *if and only if* $\alpha_i + \overline{\alpha}_i(\mu_j) \geqslant \alpha_{i'} + \overline{\alpha}_i(\mu_{j'})$.

Notice that point (1) combined with equation (1.6) reveals the "product externality" effect. To see this, suppose the measure of buyers buying products l and h from seller M are given by (μ_l, μ_h). Now suppose we rearrange the market shares of each product for some $\varepsilon > 0$ such that the new measures become $(\mu_l + \varepsilon, \mu_h - \varepsilon)$. In this case point (1) and equation (1.6) imply that the two-period utility from buying l in period 0 increases for all buyer types while the two-period utility from buying h decreases at the same rate. Hence, the efficient solution consists of finding the balance between these two effects.

Point (2) helps us determine when having an unbalanced distribution is better than a balanced distribution. As the distribution becomes unbalanced, the utility of one group of buyers increases at the expense of the other group. Point (2) implies that the gain from an unbalanced distribution is higher when information is more valuable. Finally point (3) reveals that the gain inflexible buyers receive from information is greater than the gain flexible buyers receive. It is this point that determines whether the efficient distribution is sorted or shuffled.

Point (3) also reveals that the two-period utility function satisfies the single-crossing property with respect to the maximal (minimal) posterior, where the single-crossing point is $\theta = 0$. The buyer of type $\theta = 0$ is indifferent between the two products and hence her two-period value function is not affected by uncertainty. Point (3) implies that, the buyers further away from the buyer of type 0 strongly prefer one product over the other in period 0 from a two-period point of view. Hence, the individual preference of a buyer becomes increasingly influential in determining the social surplus as the buyer's type moves further away from 0. Therefore, knowing the preference ranking of each buyer over the two products is essential to finding the right balance between the individual buyers' interests and the society's interest, which determines the efficient solution.

Point (4) implies that the preference rankings for all buyers with $\theta \neq 0$ is the same as the rankings of the sum of first period priors and the respective expected maximal (minimal) posteriors. We can divide the (μ_l, μ_h) space into regions in accordance with the different rankings of utility. Let $U = \{(\mu_l, \mu_h) \mid 0 \leqslant \mu_h, \mu_l \leqslant 1, \mu_h + \mu_l \leqslant 1\}$ represent the measure space. Let $I^+ = \left\{(\mu_l, \mu_h) \in U \mid \alpha_i + \overline{\alpha}_i\left(\mu_j\right) \geqslant \alpha_j + \bar{\alpha}_j\left(\mu_i\right) \right\}$ and $I^- = \left\{(\mu_l, \mu_h) \in U \mid \alpha_i + \underline{\alpha}_i\left(\mu_j\right) \leqslant \alpha_j + \underline{\alpha}_j\left(\mu_i\right) \right\}$ where $i, j \in \{l, h\}$ and $j \neq i$, $I = i$. In particular I^+ defines the region where buyers with positive types prefer i over j and I^- defines the same region for buyers with negative types. The following lemma which is derived directly from equation (1.7) describes these regions.

Lemma 1.2 (REGIONS THAT DETERMINE RANKINGS).

The regions are identified as follows:

1. $H^+ = \left\{(\mu_l, \mu_h) \in U \mid \mu_h - \mu_l \leqslant \frac{2}{\rho-1}\right\}$ *and* $L^+ = \left\{(\mu_l, \mu_h) \in U \mid \mu_h - \mu_l \geqslant \frac{2}{\rho-1}\right\}$

2. $L^- = \left\{(\mu_l, \mu_h) \in U \mid \mu_l - \mu_h \leqslant \frac{2}{\rho-1}\right\}$ *and* $H^- = \left\{(\mu_l, \mu_h) \in U \mid \mu_l - \mu_h \geqslant \frac{2}{\rho-1}\right\}$

Notice that we can restrict these regions even further for efficiency purposes by looking at the subsets with $\mu_h + \mu_l = 1$, because as we mentioned above we are only interested in full market share distributions. When looking for the efficient distribution (μ_l, μ_h), we have to be aware that the ranking of the utilities from the two products might change as we change (μ_l, μ_h). It is worth explaining what these regions imply. First, these regions are not disjoint as illustrated by the following figure.

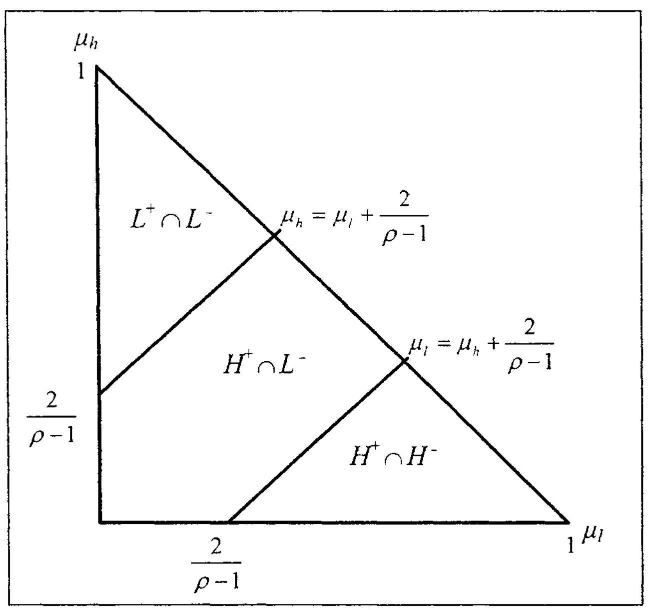

Figure 1.4: Regions of different product rankings

If $(\mu_l, \mu_h) \in (H^+ \cap L^-)$, then buyers rank the utilities in accordance with the static preferences, i.e. one-period preferences. Lemma 1.2 implies that for $(\mu_l, \mu_h) \in (H^+ \cap L^-)$, it has to be the case that $|\mu_h - \mu_l| \leqslant \frac{2}{\rho-1}$. The intuition behind this is that when the difference between the measures of buyers experiencing each product is not high, period 1 informational returns from both products are similar and hence each buyer ranks the products with respect to their period 0 returns. If $(\mu_l, \mu_h) \in L^+ \cap L^-$, then μ_h is so high that even the buyers who would prefer product h in a static world, i.e. the buyers with $\theta > 0$, get a higher two-period utility from product l than product h. The intuition is that a buyer of type $\theta > 0$ finds it worthwhile to delay her purchase of product h until the second period because there is a high probability that she will receive information on h when μ_h is high. In period 0 she buys product l from seller M to have access to the expected recommendation in the second period. Thus when $(\mu_l, \mu_h) \in L^+ \cap L^-$, all buyers get the highest two period utility from version l and similarly when $(\mu_l, \mu_h) \in H^+ \cap H^-$, all buyers get the highest two period utility from version h.

It is clear that once we restrict attention to full market share distributions, the only balanced distribution is $\left(\frac{1}{2}, \frac{1}{2}\right) \in H^+ \cap L^-$, where the rankings are in accordance with the static preferences. However, both $H^+ \cap L^-$ and $L^+ \cap L^-$ have unbalanced distributions with $\mu_l < \mu_h$. Thus, when looking for an optimal unbalanced distribution, we have to take into account the possibility that the optimal solution may be changing regions.

There is also one thing to pay attention when we look for an optimal shuffled distribution. Recall from Definition 1.2 that if a distribution with $\mu_i \geqslant \mu_j$ for some $i \in \{l, h\}$ is shuffled, then the buyers of j consist of two segments S^- and S^+ which are respectively at the left and right ends of $[-1, 1]$. The next lemma compares S^- and S^+.

Lemma 1.3. *If a shuffled distribution with* (μ_h, μ_l) *is efficient, then* $|S^-| \lesseqgtr |S^+|$ *as* $\mu_h \lesseqgtr \mu_l$.

Proof. Suppose $\mu_h \geqslant \mu_l$ and $|S^-| < |S^+|$. From equations (6) and (7) it is easily seen that $U_M(-\theta, l, \mu_h) > U_M(\theta, l, \mu_h)$ for all $\theta > 0$ and $\mu_h \geqslant 0$. But then the social surplus can be increased by swapping the products purchased by some negative and positive types without changing (μ_h, μ_l). ∎

This lemma simply means that if it is efficient for some positive types to purchase product l, then it has to be the case that there are more negative types purchasing l. The intuition is that regardless of the magnitude of μ_h, a buyer with negative type derives a weakly higher utility from l than a buyer with symmetric positive type. Hence, from now on, when we say "shuffled", we will refer to a shuffled distribution with the property in Lemma 1.3.

Proposition 1.1 (EFFICIENCY).

The efficient allocation is such that seller M *has full market share and there exists a unique* $\rho_s > 3$ *such that;*

1. *for* $\rho \leqslant \rho_s$, *the unique efficient distribution of buyers is balanced and sorted,*
2. *for* $\rho_s < \rho < \infty$, *there are two symmetric efficient distributions of buyers that are unbalanced and imperfectly shuffled.*

Proof. The proofs of all propositions are relegated to the Appendix. ∎

Figure 1.5 illustrates the efficient distributions in Proposition 1.1. Recall that a high ρ indicates either high informativeness or a low initial information, both of which increase

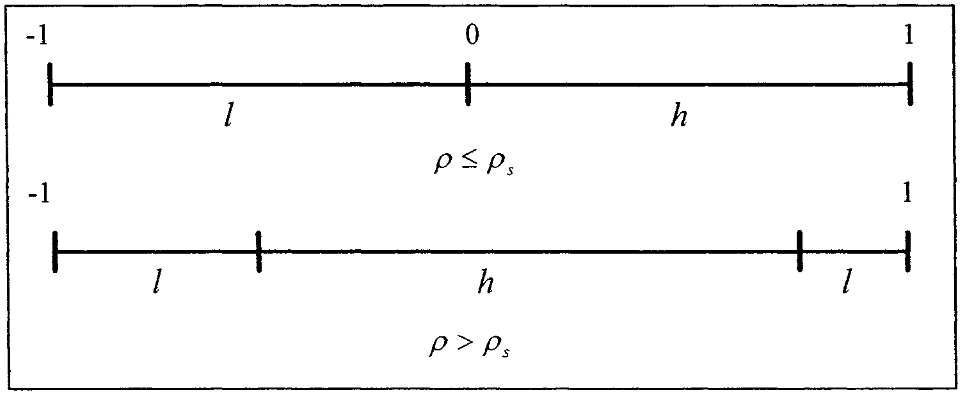

Figure 1.5: The efficient allocations with $\mu_h \geqslant \mu_l$ for $\rho \leqslant \rho_s$ and $\rho > \rho_s$ respectively

the value of the additional information that the recommender system provides. Thus, the proposition says that for low levels of ρ, the efficient allocation is no different than the efficient allocation with no recommender system because the informational contribution of the recommender system is not large enough to force the efficient allocation away from static preferences. For ρ large enough however, the information provided by the recommender system becomes valuable enough to make the efficient distribution unbalanced.

The first thing that draws attention in the proposition is the fact that $\rho_s > 3$. Lemma 1.2 implies that if $\rho < 3$, $L^+ = H^- = \emptyset$, hence the ranking of the products are in line with static preferences for all (μ_h, μ_l). First notice that if the static preferences hold, the efficient distribution is sorted. This results directly from supermodularity, which implies that inflexible buyers' preference is stronger than the flexible buyers'. Next notice that if the efficient distribution is sorted then it is balanced. To understand this, consider an unbalanced and sorted distribution with $\mu_h > \mu_l$, where there is a buyer $\widehat{\theta} < 0$ such that all buyers with type $\theta \geqslant \widehat{\theta}$ purchase product h and the remaining purchase product l. Point (1) in Lemma 1.1 implies that the utilities of buyers of h increase and the utilities of buyers of

l decrease in $\widehat{\theta}$. Moreover buyers of different products symmetric around 0 face changes of identical magnitude in opposite directions[4]. Since $\mu_h > \mu_l$, the gain of buyers of h not only covers for the loss of others but also creates an additional increase in the surplus. Hence μ_l should increase up to $\frac{1}{2}$.

For $\rho \geqslant 3$, all the preference regions are non-empty. Hence there exists (μ_h, μ_l) such that all buyers prefer h or l. Suppose we restrict attention to $(\mu_h, \mu_l) \in L^+ \cap L^-$, which implies that there is sufficient difference between μ_h and μ_l and hence a sufficient level of unbalance. For high enough ρ, the really inflexible positive types prefer l more strongly than sufficiently flexible negative types due to the high informational returns in the second period. Therefore, the optimal distribution awards them their preferred product at the expense of some flexible buyers at either side of 0. This implies a shuffled distribution. The next step is to compare the optimal shuffled distribution with the balanced and sorted distribution. The proposition implies that it is only when $\rho \geqslant \rho_s$ that the contribution of the recommender system is large enough to make it worthwhile creating unbalance through the shuffled distribution.

The next proposition shows how these dynamics change as ρ increases. Recall Proposition 1.1 shows that for $\rho < \rho_s$, the distribution stays balanced and sorted and nothing changes. The interesting dynamics are when the efficient distribution is unbalanced and shuffled.

Proposition 1.2 (COMPARATIVE STATIC FOR EFFICIENCY).

[4]Notice that these effects are identical becuase the probability of receiving a signal on i is linear in μ_i. However, notice that linearity is sufficient but not necessary for the result that a sorted distribution has to be balanced. Any function that is concave in μ_i would generate the same result. In fact all our results would easily go through with a probability function that is concave in μ_i.

1. *At $\rho = \rho_s$, the degree of unbalance and the degree of shuffling increase discontinuously.*

2. *For $\rho \geqslant \rho_s$, the degree of unbalance and shuffling increase in ρ.*

3. *As $\rho \to \infty$, the distribution for both efficient allocations becomes perfectly shuffled, where all buyers with type $\theta \in \left[-\frac{1+\sqrt{7}}{6}, \frac{1+\sqrt{7}}{6}\right]$ buy product i and all the others buy product j for some $i \in \{l, h\}$.*

Proposition 1.2 says that as information becomes more valuable it is beneficial to increase the degree of unbalance and place a higher burden on flexible buyers. As there is more that the recommender system can contribute, the inflexible buyers increasingly have more to gain than flexible buyers. Thus the total surplus increases if the inflexible buyers are satisfied at the expense of flexible buyers. The interesting feature is that the transition from the sorted distribution to the shuffled distribution is discontinuous. The reason is that as explained above an efficient distribution can be unbalanced only if it is shuffled. A shuffled distribution implies that the inflexible buyers of both types purchase the same product $i \in \{l, h\}$ and the flexible buyers purchase product $j \neq i$. But due to supermodularity, this can not be efficient if there is an inflexible buyer who prefers j, because then we could always find a flexible buyer and swap their purchases. Hence it must be the case that all buyers prefer the same product i in period 0. Lemma 1.2 shows, however, that all buyers prefer product i in period 0 if and only if there is sufficiently more buyers buying version i than version j. The necessity for such a discrete difference between the two measures brings about the discontinuous jump.

1.5 Equilibrium

In this section we investigate Perfect Bayesian Equilibria of the game between seller M and the buyers. Equilibrium will be determined by seller M's incentives, whereas the efficient allocation was determined by social incentives. The socially optimal allocation internalizes all the informational externality inherent in the problem. However, as we will see since the seller cannot employ a first degree price discrimination, he is not able to internalize the externality to the full extent. This creates a discrepancy between the equilibrium distributions and the efficient one.

As we described in Section 1.2, seller M announces a price for each product in period 0 and each buyer optimally chooses a seller and a product given the prices. Similar to the efficiency case, finding the equilibrium in this setting is very straight forward if there is no recommender system, which is equivalent to $\varepsilon = \frac{1}{2}$ or $\gamma = 0$. In either of these cases, seller M and fringe are effectively selling identical products. Thus the competition is fierce, which implies Bertrand solution for each period and therefore coincides with the efficient allocation. When we introduce some uncertainty and informativeness into the setting, as before, period 0 distribution of buyers affects the information gathered for period 1 and hence the utilities in period 1. Here seller M has the sole control over the distribution of buyers through the prices he sets in period 0. In other words seller M is responsible for how much information is gathered for period 1. In period 1, seller M reveals the information he has gathered to the buyers who have bought from him in period 0 and announces new prices. The fact that the seller can make the information distribution conditional on period 0 purchases allows him to charge the buyers in period 0 for the information they will receive in period 1. Therefore, he may extract some of the informational benefits through higher prices. This gives him incentives to choose his pricing scheme in a way to collect sufficient amounts of information. The extent to which he can extract these benefits determines how similar the distribution he creates is to the efficient one.

We will start analyzing seller M's problem with the subgame in period 1.

Lemma 1.4 (SECOND PERIOD SUBGAME).

The minimum price in the market in period 1 in any perfect Bayesian equilibrium equals marginal cost for each product.

Proof. In the second period, seller M and seller F are selling the exact same products and supplying the same service. Thus the problem is identical to the Bertrand problem ∎

This lemma is due to the fact that a buyer can get the recommendation from seller M and yet purchase from the fringe given the recommendation. The services all sellers provide are identical in period 1 and the competition is at the Bertrand level. The interesting part of the problem is period 0 prices.

We first examine the subgame played between the buyers after seller M announces $\mathbf{p} = (p_h, p_l)$. Here, we need to start taking the fringe into account, because seller M's pricing affects the seller choice of a buyer as well as the product choice. Although efficiency requires seller M to have full market share, there is no guarantee that seller M will choose to have full market share. We introduce the following definition regarding this issue.

Definition 1.3 (GAP). *For a given distribution of buyers with (μ_i, μ_j) where $\mu_i \geqslant \mu_j$ for some $i \in \{l, h\}$, if the distribution satisfies the criteria for one of the definitions given in*

Definition 1.2 except that a segment around zero separates either the buyers of j into two segments or separates the buyers of i and j, we say it satisfies the definition "with a gap".

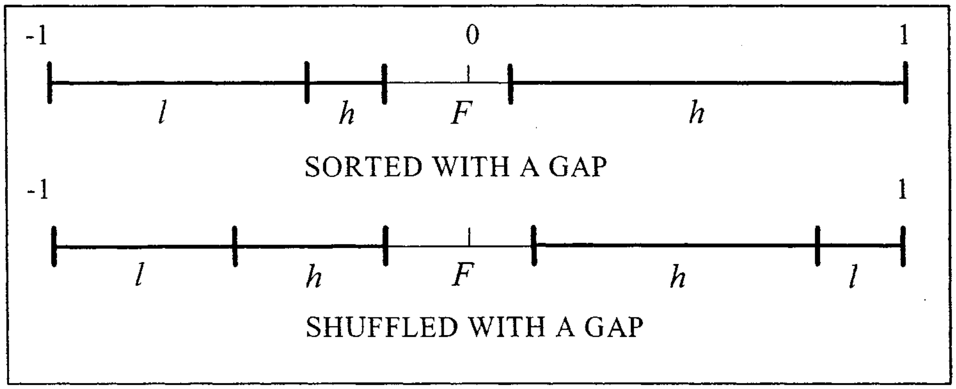

Figure 1.6: The Distributions with a Gap

To determine a buyer's seller choice, we need to know the utility she gets when she purchases either product from the fringe. Notice that if a buyer purchases product i from the fringe in period 0, she will maximize her expected utility choosing from $\{j, m\}$ in the second period without any additional information. Then, the expected maximal (minimal) posterior in equation (1.7) with $\mu_l = \mu_h = 0$ determines her expected utility in period 1. Hence, the relevant variables are $\overline{\alpha}_i(0)$ and $\underline{\alpha}_i(0)$ and the two-period value function for a buyer of type θ conditional on purchasing product i from seller F in period 0 is

$$U_F(\theta, i) = U_M(\theta, i, 0) \tag{1.9}$$

Given this interpretation of the utilities, the properties given by Lemma 1.1 applies to the fringe utilities as well. Point (1) in Lemma 1.1 combined with point (4) implies that buying a product from seller M always gives weakly higher two-period gross utility than buying the same product from the fringe. The following lemma is derived directly from equation (1.7) and gives some further hints towards a buyer's product ranking when purchasing from the fringe.

Lemma 1.5 (FRINGE UTILITIES). *For all $i \in \{l, h\}$ and all $\mu_j \geqslant 0$, $\alpha_h + \bar{\alpha}_h(0) \geqslant \alpha_l + \bar{\alpha}_l(0)$ and $\alpha_l + \underline{\alpha}_l(0) \leqslant \alpha_h + \underline{\alpha}_h(0)$.*

This lemma and point (4) in Lemma 1.1 together imply that, if purchasing from the fringe, buyers of type $\theta \geqslant 0$ choose h and buyers of type $\theta \leqslant 0$ choose l. Thus, in the remainder of the section F will refer to these optimal choices for each group of buyer. Supermodularity holds for the fringe value function as well, where the single crossing point is still $\theta = 0$. The next important question to answer is how the fringe utility fits into the utility rankings. For that we introduce the following regions: let

$$\begin{aligned} LF^+ &= \{(\mu_l, \mu_h) \in U \mid \alpha_l + \bar{\alpha}_l(\mu_h) \geqslant \alpha_h + \bar{\alpha}_h(0)\} \\ HF^- &= \left\{(\mu_l, \mu_h) \in U \mid \alpha_h + \underline{\alpha}_h(\mu_j) \leqslant \alpha_l + \underline{\alpha}_l(0)\right\} \end{aligned}$$

and FL^+ and FH^- are defined as the complements respectively. For $I \in \{l, h\}$, IF^+ refers to the region where the buyers with positive types prefer purchasing $i = I$ from seller M to purchasing h from the fringe and FI^+ is the region where the preference is reversed. IF^- and FI^+ define the similar regions for negative types. The following lemma identifies these regions:

Lemma 1.6 (REGIONS THAT DETERMINE FRINGE RANKING).

1. $LF^+ = \left\{(\mu_l, \mu_h) \in U \mid \mu_h \geqslant \frac{2}{\rho-1}\right\}$ and $FL^+ = \left\{(\mu_l, \mu_h) \in U \mid \mu_h \leqslant \frac{2}{\rho-1}\right\}$

2. $HF^{-}=\left\{(\mu_l,\mu_h)\in U\mid \mu_l\geqslant\frac{2}{\rho-1}\right\}$ *and* $FH^{-}=\left\{(\mu_l,\mu_h)\in U\mid \mu_l\leqslant\frac{2}{\rho-1}\right\}$

Region LF^{+} is the region where buyers of type $\theta>0$ prefer purchasing product l from seller M to purchasing product h from the fringe. Notice that for this to occur, μ_h has to be sufficiently high, because the buyer is giving up her static preference for informational returns. Region FL^{+} defines the complementary region. HF^{-} and FH^{+} are defined similarly. These regions, together with the regions defined in the previous section reveal how the buyers rank the two-period utilities from different products and sellers depending on (μ_l,μ_h). The buyers however, do not observe (μ_l,μ_h) directly, but form a belief about it after observing the prices. Thus, the next step is to see what a buyer's optimal choice is, given the prices and her belief about other's actions, which determines her utility ranking.

Let $(\widehat{\mu}_l,\widehat{\mu}_h)$ represent the belief of a buyer about others actions after observing $\mathbf{p}$. Let

$$U\left(\theta,i,s,\widehat{\mu}_j,p_i\right)=\begin{cases}U_M\left(\theta,i,\widehat{\mu}_j\right)-p_i-c \text{ if } s=M\\ U_M\left(\theta,i,0\right)-2c \text{ if } s=F\end{cases}\tag{1.10}$$

Each buyer will maximize $U\left(\theta,i,s,\widehat{\mu}_j,p_i\right)$ with respect to i and s given her belief $(\widehat{\mu}_l,\widehat{\mu}_h)$. These choices will generate a measure vector (μ_l,μ_h). The equilibrium occurs when the beliefs coincide with the measures generated by the optimal choices. Due to the coordination element inherent in the problem multiple equilibria might arise. We leave the discussion about the multiple equilibria to the appendix, where we show that the equilibria in the subgame following the price announcement are pareto ranked. Moreover an increase in pareto rank coincides with an increase in the vector (μ_l,μ_h), which hence implies an increase in the seller's profits. If we consider the two products as two networks, then following the literature and in particular Economides and Himmelberg (1995) and Economides (1996) we can expect the large network sizes to arise. Hence from this point on we will choose the Pareto optimal equilibrium of this subgame. Let $(\mu_h(\mathbf{p}), \mu_l(\mathbf{p}))$ represent the market share of seller M for each product in this equilibrium when he announces prices $\mathbf{p}$.

So far, the focus has been on the market shares that result from the announcement of a price vector. For the seller, however, the buyer composition of the market shares that the price vector generates is also important. Due to the single-crossing property that the two-period value functions employ, after the announcement of any price vector, the market is segmented into finitely many segments of buyers, where each segment is identified by a unique choice of product and seller. The buyer types that separate these segments are the "marginal buyers", who are indifferent between the choices of the two neighboring segments. The purchasing choices of each segment can be expressed in terms of some incentive constraints and the marginal buyers are the buyers for whom the incentive constraints hold with equality. Recall the regions into which we divided the (μ_h, μ_l) space. Those regions represent different rankings of utilities between different product and seller choices. Thus in the context of seller M's pricing choice, they imply different marginal buyer structures, i.e. different kinds of segmentation. This is important, because as the seller changes the pricing, $(\mu_l(\mathbf{p}), \mu_h(\mathbf{p}))$ changes and so does the segmentation.

We can write the profit of seller M as a function of $\mathbf{p}$, knowing that it will generate $(\mu_l(\mathbf{p}), \mu_h(\mathbf{p}))$. Alternatively we can take a dual approach and write the profits as a function of the market shares (μ_h, μ_l), which imply a particular price vector $(p_h(\boldsymbol{\mu}), p_l(\boldsymbol{\mu}))$ that generates them[5]. In this paper, we take the dual approach of choosing the market

[5]Notice that $\boldsymbol{\mu}(\mathbf{p}) = (\mu_l(\mathbf{p}), \mu_h(\mathbf{p}))$ is a one-to-one function. Suppose that there are two vectors $\mathbf{p}$ and $\mathbf{p}'$ such that $\boldsymbol{\mu}(\mathbf{p}) = \boldsymbol{\mu}(\mathbf{p}')$ and $\mu_l(\mathbf{p}) + \mu_h(\mathbf{p}) < 1$. Given the differential prices and the fact that $\boldsymbol{\mu}(\mathbf{p}) = \boldsymbol{\mu}(\mathbf{p}')$, there will be at least one product i for which there will less buyers willing to purchase it

shares. Because when market shares are the main variables, the regional changes and thus the changes in segmentation is directly observable. We can write the profits as a function of (μ_h, μ_l) as

$$\pi_M(\mu_h, \mu_l) = \mu_h(p_h(\boldsymbol{\mu}) - c) + \mu_l(p_l(\boldsymbol{\mu}) - c) \tag{1.11}$$

Seller M chooses (μ_h, μ_l) to maximize these profits. Let us look at the trade-off involved in increasing μ_h:

$$\frac{\partial \pi_M(\mu_h, \mu_l)}{\partial \mu_h} = \left[\mu_h \underset{(-)}{\frac{\partial p_h(\boldsymbol{\mu})}{\partial \mu_h}}\right] + \left[(p_h(\boldsymbol{\mu}) - c) + \mu_l \underset{(+)}{\frac{\partial p_l(\boldsymbol{\mu})}{\partial \mu_h}}\right] \tag{1.12}$$

The first bracket represents the marginal loss, i.e. the decrease in p_h that is required to increase μ_h. This is the typical loss a monopolist incurs in a standard profit maximization problem. The first term in the second bracket represents the direct marginal gain, i.e. the fact that the mark-up is received from a higher market share. Again this is the gain we would see in a standard monopolist problem. What makes this problem different is the last term in the second bracket, which is a direct reflection of the product externality. The last term represents the indirect marginal gain which is due to the fact that as μ_h increases the two-period utility from l increases, hence the same μ_l could be kept at a higher p_l. The trade-off between the loss and gain determines the optimal market shares for the seller. The following propositions give the solution to all these effects and reveal the equilibria.

Proposition 1.3 (EQUILIBRIUM 1).

---

under one price vector than the other. But then they cannot generate the same measure of buyers. Consider the same situation for when $\mu_l(\mathbf{p}) + \mu_h(\mathbf{p}) = 1$. The same arguement holds up to prices below c, which we can simply exclude. Hence $(p_h(\boldsymbol{\mu}), p_l(\boldsymbol{\mu}))$ is the inverse of $(\mu_l(\mathbf{p}), \mu_h(\mathbf{p}))$.

There exists a unique $0 < \rho_1 < \infty$ *such that for* $\rho \leqslant \rho_1$ *the equilibrium is unique and is characterized by*

1. *prices* $p_i^* = \frac{1}{18} \max \left\{ \frac{1}{2}\gamma \left(1 - 4\varepsilon^2\right), (\gamma - \varepsilon) \right\} + c$ *for all* $i \in \{l, h\}$,
2. *less than full market share for seller* M,
3. *a balanced and sorted distribution with a gap where the set of buyers buying products* l *and* h *from seller* M *are respectively* $\left[-1, -\frac{1}{3}\right]$ *and* $\left[\frac{1}{3}, 1\right]$.

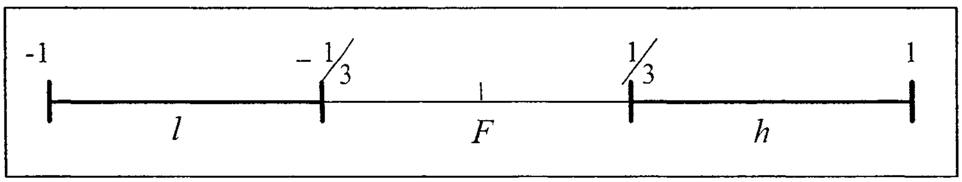

Figure 1.7: The equilibrium for $\rho \leqslant \rho_1$

Proposition 1.3 gives the unique equilibrium for low levels of ρ. Low ρ means either γ is low or that ε is high. In either case, the recommender system does not play a big role in reducing uncertainty. The first thing to understand is why the seller prefers less than full market share in this case. It is clear that it can not be optimal to have full market share and a balanced distribution, because it yields zero profits and the seller certainly has other options giving him strictly positive profits. In consequence, the optimality of full market share necessitates an unbalanced distribution. Increasing the market share has the cost and benefits, which were discussed in marginal terms in equation (1.12). The direct gain is that more buyers purchase from seller M. The direct loss is that it requires an initial prices decrease for at least one product. However, this loss is dampened because the increase in the market share leads to an increase in the utility from buying some product from seller M. This is a result of the product externality we described earlier. Therefore, as an indirect gain, the seller will either be able to not decrease the price as much to generate the same market share increase or increase the price of one product while decreasing the other. Consider the two-period utilities normalized by ε. When ρ is low, equations (1.6) and (1.7) imply that the utility difference between purchasing from sellers M and F does not decrease by much as the buyer's type gets more flexible. Hence, the initial price decrease needed to generate a given market share increase is not large. However, the indirect gain is not large either. Because, by Lemma 1.1, if ρ is low, the normalized utility of a buyer increases by very little as market share increases. Proposition 1.4 says that for low ρ, the indirect gain is not strong enough compared to the direct price decrease effect and thus the seller chooses to leave out some buyers. The similar reasoning applies to the choice of degree of balance. Therefore, for low ρ, the seller prefers to make profits simply by increasing the price equally on both products to a level that sufficiently inflexible buyers are willing to pay to have access to new information in period 1. The seller's problem can be interpreted as separated into two disjoint markets, in each of which he sells a higher quality product compared to the fringe and thus sets a higher price.

The next propositions show that as ρ increases, the seller finds it optimal to treat his problem as a more complex problem than the analogy of the "high quality problem" and prefers to explore the different informational structures he can attain through differential pricing.

Proposition 1.4 (EQUILIBRIUM 2).

There exists $\rho_3 > \rho_2 > \rho_1$ such that for $\rho_1 < \rho \leqslant \rho_3$ there exists two symmetric equilibria *identified with $i \in \{l, h\}$ and characterized by*

1. $p_i^* > c$ *and* $p_j^* \begin{cases} < c \text{ if } \rho_1 < \rho < \rho_2 \\ = c \text{ if } \rho_2 \leqslant \rho \leqslant \rho_3 \end{cases}$

2. *full market share for seller M,*

3. *an unbalanced and sorted distribution where $\mu_i^* < \mu_j^*$.*

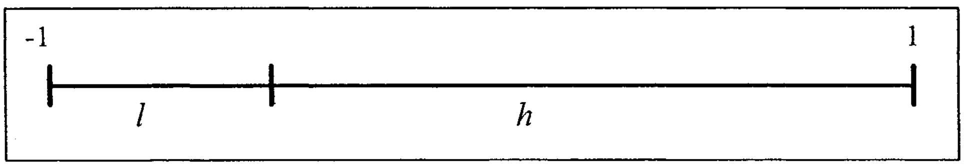

Figure 1.8: One of the two symmetric equilibria for $\rho_1 < \rho \leqslant \rho_3$

Proposition 1.4 first reveals that there exists intermediate values of ρ for which the seller captures the whole market. As explained above, the direct loss due to increasing the market share and the degree of unbalance, i.e. the direct price decrease, increases with ρ, because the buyers get more differentiated with respect to how much they prefer buying from seller M to the fringe. However, the indirect gain also increases in ρ since the utilities become more responsive to changes in the market shares. Propositions 1.3 and 1.4 say that the indirect gain increases faster than the direct loss. The more intriguing thing is that for $\rho_1 < \rho < \rho_2$, the seller is willing to make a loss on one product, because this allows a price increase on the other product that more than covers the loss. Then as ρ increases over ρ_2, even for the buyers with type $\theta < 0$, buying product h from seller M becomes a better choice than buying product l from seller F and thus the necessity to decrease the price below marginal cost disappears.

Proposition 1.5 (EQUILIBRIUM 3).

There exists a $\rho_4 > \rho_3$ such that for $\rho_3 < \rho \leqslant \rho_4$ there exists two symmetric equilibria identified with $i \in \{l, h\}$ and characterized by

1. *prices $p_i^* > c$ and $p_j^* = c$,*
2. *full market share for seller M,*
3. *an unbalanced and shuffled distribution where $\mu_i^* < \mu_j^*$.*

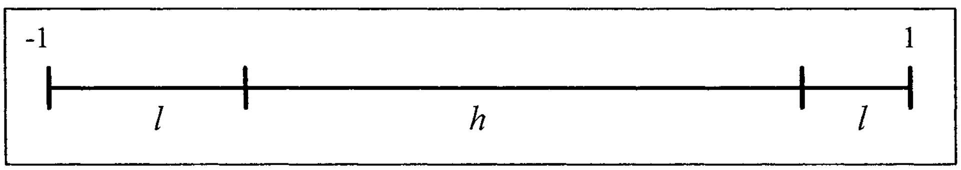

Figure 1.9: One of the two symmetric equilibria for $\rho_3 < \rho \leqslant \rho_4$

Proposition 1.5 shows that until $\rho = \rho_4$, the indirect gain dominates the direct loss. Moreover, ρ here is so high that even the buyers of type $\theta > 0$ prefer product l to product h when buying from seller M. The seller makes use of this preference structure by including the inflexible buyers of both types in the group that pays a high price to receive the information provided by the flexible buyers. Therefore the distribution becomes shuffled.

Proposition 1.6 (EQUILIBRIUM 4).

For $\rho > \rho_4$ there exists two symmetric equilibria identified with $i \in \{l, h\}$ and characterized by

1. *prices $p_i^* (\gamma, \varepsilon) > p_j^* (\gamma, \varepsilon) > c$,*
2. *less than full market share for seller M,*

3. *an unbalanced and shuffled distribution with a gap where* $\mu_i^* < \mu_j^*$.

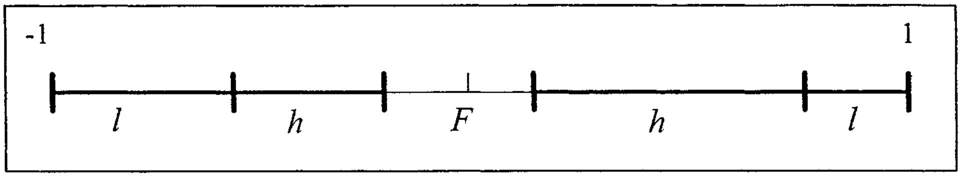

Figure 1.10: One of the two symmetric equilibria for $\rho > \rho_4$

After $\rho = \rho_4$, we see that the pattern is reversed. The seller chooses a less than full market share, because once again, the indirect gain from increasing the market share becomes smaller than the direct loss. For $\rho > \rho_4$, the utility difference between buying l from seller M and buying from F decreases sharply as the type gets more flexible. Hence, the price increase due to leaving out some buyers of one product is so high that it more than compensates for the loss incurred on the buyers of the other product. As the next proposition shows the reversal of these two effects also implies that the total market share and the degree of unbalance keep decreasing for all $\rho > \rho_4$.

Propositions 1.3-1.6 have shown that the seller utilizes recommender system the most for intermediate levels of ρ. Because, for low levels of ρ informational effects are not high enough to make it profitable to capture the whole market and for high levels of ρ, the seller can make large gains through very high prices by leaving out a small measure of buyers . Proposition 1.7 shows the comparative statics these effects generate.

Proposition 1.7 (COMPARATIVE STATIC).

In any subgame perfect equilibria ,

1. *if* $\rho < \rho_1$ *the measures of buyers buying either product from either seller do not change*

with ρ,

2. *if* $\rho_1 \leqslant \rho < \rho_3$*, the degree of unbalance increases in* ρ*,*

3. *if* $\rho_3 \leqslant \rho < \rho_4$*, the degree of unbalance decreases and the degree of shuffling increases in* ρ

4. *if* $\rho \geqslant \rho_4$*, total market share and the degree of unbalance decreases in* ρ*,*

5. *as* $\rho \to \infty$*, the distribution of buyers becomes perfectly shuffled where the set of buyers buying products* i *and* j *are* $\left[-1, -\frac{7}{9}\right] \cup \left[\frac{7}{9}, 1\right]$ *and* $\left[-\frac{7}{9}, -\frac{1}{9}\right] \cup \left[\frac{1}{9}, \frac{7}{9}\right]$ *respectively for some* $i \in \{l, h\}$*.*

1.6 Welfare

In this section we investigate how the efficient allocation compares to the equilibrium allocation. The market shares optimal for seller M may not be optimal for the society. The intuition is that, when considering the effect of a policy change, the seller internalizes the effect in only the marginal buyers' utilities. For the society, however, the change in the utilities of all the buyers matters. The first very intuitive discrepancy this may create between the equilibrium and the efficient allocation is that, full market share is not necessarily optimal for seller M. He may find it too costly to capture the whole market, because he does not internalize all the gains from an increased market share. The other discrepancies may be in the level of balance and shuffling the seller chooses. The next figure illustrates the properties of the efficient allocation and the equilibrium allocation for different levels of ρ.

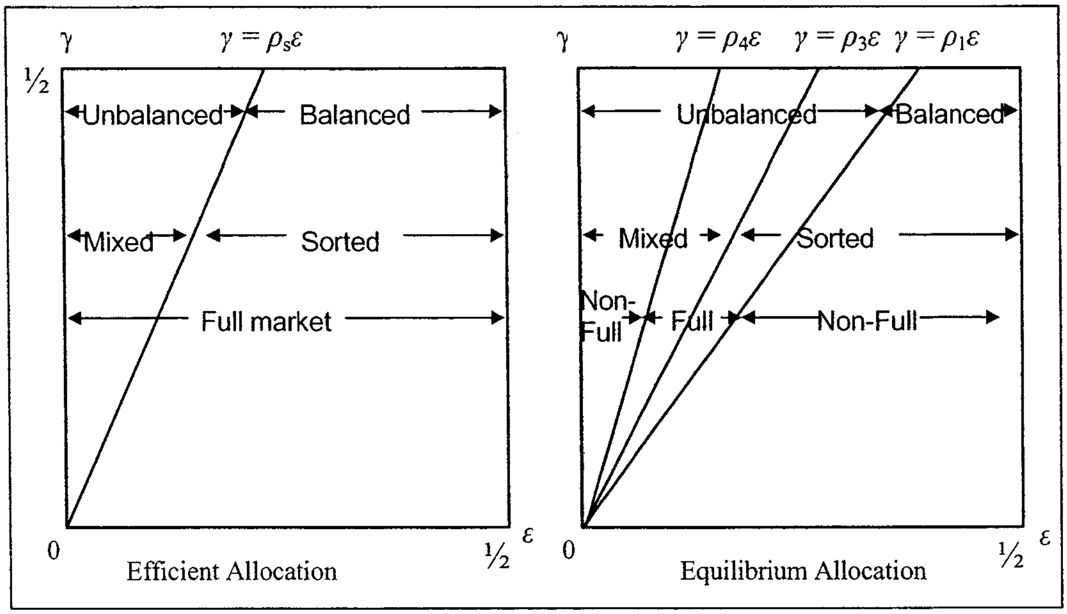

Figure 1.11: Equilibrium and efficient allocations as a function of γ and ε

The first difference, which is clearly seen in Figure 1.11, is that for $\rho \leqslant \rho_1$ and $\rho \geqslant \rho_4$, the equilibrium leads to under-utilization of the recommender system because it leaves some buyers to the fringe. The second important difference is regarding the degree of unbalance.

Proposition 1.8 (EQUILIBRIUM VS EFFICIENCY).

1. *For all $\rho \leqslant \rho_1$ both the equilibrium and the efficient distributions are balanced.*

2. *For all $\rho > \rho_1$, the degree of unbalance is higher for each equilibrium distribution than the corresponding efficient distribution.*

3. *The difference between the degree of unbalance increases in ρ if $\rho_1 < \rho \leqslant \rho_3$ and decreases in ρ if $\rho > \rho_3$.*

The last proposition reveals another difference between the equilibrium allocations and the efficient allocations. Even when the seller utilizes the recommender system to the full extent, the system is over-utilized for one product and under-utilized for the other. Notice that for very low ρ and very high ρ the main source of inefficiency comes from the fact that the seller does not capture the whole market, whereas for intermediate levels of ρ the source of inefficiency is the discrepancy in the degree of balance. The seller increases his share of the market for one product beyond the optimal level because he does not internalize the loss in the utilities of all the buyers of that product, but he only internalizes the loss in the utility of the marginal buyer through lower prices.

The next question we would like to ask is whether the efficient allocation can be decentralized. Because then we can compare the decentralizing prices with the equilibrium prices and discover how the discrepancy between efficiency and equilibrium is created. For the purpose of the next proposition let (μ_h^s, μ_l^s) denote the efficient measures of buyers purchasing each product.

Proposition 1.9 (DECENTRALIZATION).

The efficient allocation with $\mu_i^s \leqslant \mu_j^s$ for some $i \in \{l, h\}$ is decentralized uniquely (up to prices below c) by

1. $(p_h^s, p_l^s) = (c, c)$ *for* $\rho \leqslant \rho_s$
2. $\left(p_i^s, p_j^s\right) = \left(U_M\left(\theta_1^s, i, \mu_j^s\right) - U_M\left(\theta_1^s, j, \mu_i^s\right) + c, c\right)$ *for* $\rho > \rho_s$, *where*

$$\theta_1^s = \begin{cases} -\mu_h^s + \frac{2\mu_h^s}{(2\mu_h - 1)(\rho - 1)} & \text{if } i = l \\ -\mu_l^s - \frac{2\mu_l^s}{\left(2\mu_l^s - 1\right)(\rho - 1)} & \text{if } i = h \end{cases}$$

The first part of the proposition is straight forward. To understand the second part let us note that the proof shows that in the efficient allocation μ_i^s is sufficiently high that

$\left(\mu_i^s, \mu_j^s\right) \in JF^+ \cap JF^-$ where $J = j$, i.e. all the buyers prefer seller M to seller F for any product. This implies that to get full market share it has to be the case that $p_j = c$. To determine p_i note that the θ_1^s given by the proposition defines the buyer with the negative type around whom the efficient allocation switches. The second switch point is $\theta_1^s + 2\mu_j^s$ and efficiency requires

$$U_M\left(\theta_1^s, i, \mu_j^s\right) - U_M\left(\theta_1^s, j, \mu_i^s\right) = U_M\left(\theta_1^s + 2\mu_j^s, i, \mu_j^s\right) - U_M\left(\theta_1^s + 2\mu_j^s, j, \mu_i^s\right)$$

In other words the two buyers identifying the switch points should have the exact same gain from purchasing i over purchasing j, otherwise we could increase the total surplus by moving θ_1^s and $\theta_1^s + 2\mu_j^s$ by the same amount to the direction of the higher gain. Notice that the switch-point buyers are required to be the marginal buyers and hence should be made indifferent between the two choices in equilibrium. But then, announcing p_i such that the price difference between i and j is equal to the amount of the gain of the switch-point buyers is both necessary and sufficient for indifference. Given these prices, by supermodularity, all the buyers with more inflexible types than the marginal buyers would purchase product i and all the flexible buyers in between the marginal buyers would purchase product j.

We can consider the products and the recommendations as a bundle. Since buyers can purchase the product in the bundle for c from the fringe, the decentralizing prices over c represents the social price of information in the bundle. The proposition says that when ρ is low, information should be provided for free, but as ρ increases the group of buyers who benefit from a higher volume of information should pay a positive premium. Now we can compare the prices that the seller charges to the decentralizing prices.

Proposition 1.10 (DECENTRALIZING VERSUS EQUILIBRIUM PRICES). *Consider the equilibrium and the efficient allocation with $\mu_i \leqslant \mu_j$ for some $i \in \{l, h\}$, then*

1. *for all $\rho < \rho_1$, $p_k^* > p_k^s$ for all $k \in \{l, h\}$;*
2. *for all $\rho_1 \leqslant \rho < \rho_2$, $p_i^* > p_i^s$ and $p_j^* < p_j^s$;*
3. *for all $\rho_2 \leqslant \rho < \rho_4$, $p_i^* > p_i^s$ and $p_j^* = p_j^s$;*
4. *for all $\rho \geqslant \rho_4$, $p_i^* - p_j^* > p_i^s - p_j^s$ and $p_j^* > p_j^s$.*

Proposition 1.10 shows that the seller prices at least one product higher than the decentralizing prices. This means he forces some buyers to pay more than the social price of information. In particular, when he captures full market, he sometimes prices one product below the efficient level to drive up the sales of that product and hence creates more information than the socially optimal level. In this case we can say that the seller is pricing information too low for the group of buyers who are the main providers of information and too high for the buyers who are mainly the users of that information. When he wastes some information by leaving out some buyers to the fringe, he prices the information too high for all buyers.

1.7 Conclusion

Through a two-period, two-product model, we have shown that the existence of a recommender system not only creates additional surplus but also introduces informational externalities into the pricing problem of the seller. If the output of the recommender system were independent of sales, then, for a seller, employing a recommender system would be equivalent to offering a high quality product in a horizontally differentiated market. This problem would be very standard and the seller would simply reflect this high quality on high prices. Our findings showed that when the recommender system does not contribute much to the market, this is still the way the seller treats the problem by segmenting the market into inflexible buyers, who agree to pay a high price for the high quality service and flexible buyers, who are left to buy elsewhere.

However, when the contribution of the recommender system increases, the seller's problem includes concerns that relate to gathering the optimal level of information on each product. We showed that in this case the seller creates endogenous differentiation between otherwise symmetric products by segmenting the market into two groups: (1) a large group of flexible buyers who constitute the experimenters and pay lower prices in return for the service they provide, (2) a smaller group of inflexible buyers who pay higher prices to have access to the feedback from the first group. The optimal segmentation for the seller is not necessarily optimal for the society. The full potential of the recommender system is not realized by the pricing scheme implemented by the seller because the seller might waste some information by not capturing the whole market. Moreover, even when he captures full market share, he chooses to over-utilize the system for some products and under-utilize it for others.

It is also important to mention that these qualitative results are not generated by the modelling choices. Our results would be valid for any specification of the signal probability that is concave in the measure of buyers. Similarly any preference structure that incorporates inflexible tastes and flexible tastes would generate similar qualitative results. Having only two-periods is restrictive, however in any finite period model, arguably the idea will be the same but it might be interesting to see how the market structure evolves with time. Therefore as a next step we would like to extend the model to a dynamic level, where the buyers enjoy the choice of making their purchases at different points in time. Then, the first input for the system, i.e. the first buyers, will be especially valuable for the seller and thus we would expect them to enjoy a premium in the form of lower prices. As more buyers make purchases and leave feedback, the seller's information becomes more valuable and thus we would expect the price of information to increase over time.

As we discussed in the introduction, there are a few things that our model does not incorporate. First, recommender system can be used to increase sales through encouraging cross-sales or turning browsers into shoppers. This kind of interpretation of recommender system is a matter of future research. It would be interesting to see to the interaction between the incentives a seller has to increase sales and/or prices using a recommender system. Second, it is possible that non-loyal customers are also asked to leave feedback about the products they have purchased from other sellers once they log onto a particular seller's website. This only enlarges the database the seller keeps on each product, enhances the quality of the service he provides and hence contributes to his further extraction of the surplus. In an environment with taste uncertainty, non-loyal customers would clearly have an incentive to reveal their feedback, because it could be used to improve their future recommendations. Providing future recommendations to non-loyal customers brings about the concept of "recommendations independent of sales". There are some internet sellers practicing this kind of recommendation mechanisms. Clearly, without the loyalty factor, he sellers can not use this practice to increase prices. However, the fact that they are willing to offer recommendations this way signals that it somehow pays off. As an extension we would like to model how the recommender system contributes to the seller's profit in a setting where non-loyalty does not exclude a customer from the recommendation mechanism. This can also be considered as the unbundling of the recommendations and the products, which creates the possibility for the seller to charge for the recommendations separately. More formally, in this case the seller can employ mixed bundling, which means that the loyal customers would get a better rate through the bundle than the non-loyal customers.

Another venue for future research for economists involves the design of recommendation mechanisms. Currently, all the research by computer scientists focuses on either writing the most efficient or predictive recommender system. However, strategic concerns are not included in the process of writing the program for a recommender system. It is clear that the recommender system is a mechanism and trying to design the most profitable mechanism for the seller would be an interesting challenge for future research.

1.8 Appendix

1.8.1 Proofs of the Propositions

Proof of Proposition 1.1. Point (2) of Lemma 1.1 shows efficiency requires full market share.

We will use total surplus as the measure of efficiency. Let $\mathbf{i} : [-1, 1] \longrightarrow \{l, h\}$ represent a possible allocation function and let $\mu_h(\mathbf{i})$ and $\mu_l(\mathbf{i})$ be the measures of buyers buying product h and l from seller M. as generated by that distribution function .The total surplus is

$$W(\mathbf{i}) = \int_{-1}^{1} \left(U_M(\theta, h, \mu_l) I_{\{\mathbf{i}(\theta)=h\}} + U_M(\theta, l, \mu_h) I_{\{\mathbf{i}(\theta)=l\}} \right) d\theta \tag{A1}$$
$$= 4v - \frac{16}{3} + 4\int_{-1}^{0} \sum_{i\in\{l,h\}} (\alpha_i + \underline{\alpha}_i(\mu_j)) I_{\{\mathbf{i}(\theta)=i\}} \theta d\theta$$
$$+4\int_{0}^{1} \sum_{i\in\{l,h\}} (\alpha_i + \overline{\alpha}_i(\mu_j)) I_{\{\mathbf{i}(\theta)=i\}} \theta d\theta$$

Let $\mathbf{i}^s(\theta)$ be the total surplus maximizing allocation and

$$\bar{i}(\boldsymbol{\mu}) \equiv \arg\max_{i\in\{l,h\}} (\alpha_i + \overline{\alpha}_i(\mu_j)) \tag{A2}$$
$$\underline{i}(\boldsymbol{\mu}) \equiv \arg\min_{i\in\{l,h\}} (\alpha_i + \underline{\alpha}_i(\mu_j))$$

Also let $\bar{j}(\boldsymbol{\mu}) \equiv \{l,h\}/\bar{i}(\boldsymbol{\mu})$ and $\underline{j}(\boldsymbol{\mu}) \equiv \{l,h\}/\underline{i}(\boldsymbol{\mu})$.

Referring to the regions we described before, we can say that

$$\bar{i}(\boldsymbol{\mu}) = \begin{cases} h \text{ if } \boldsymbol{\mu} \in H^+ \\ l \text{ otherwise} \end{cases} \quad \text{and } \underline{i}(\boldsymbol{\mu}) = \begin{cases} l \text{ if } \boldsymbol{\mu} \in L^- \\ h \text{ otherwise} \end{cases}$$

Given these definitions, the proof consists of the following steps:

(1) First we will show that the distribution that $\mathbf{i}^s(\theta)$ implies has to be of the form

$$\overset{\underline{i}(\mu)}{[-1,\theta_1]}, \overset{\underline{j}(\mu)}{[\theta_1,0]}, \overset{\bar{j}(\mu)}{[0,\theta_2]}, \overset{\bar{i}(\mu)}{[\theta_2,1]} \tag{A3}$$

where $-1 \leqslant \theta_1 \leqslant 0 \leqslant \theta_2 \leqslant 1$ and the labels represent what the buyers in the corresponding interval should purchase in period 0. Suppose (WLOG) that there exists two points $0 < \underline{\theta} < \overline{\theta} < 1$ and a $v > 0$ such that the distribution generated by the optimal allocation $\mathbf{i}^s(\theta)$

looks like the following for $\theta \geq 0$:

$$..., [0, .], ..., \overset{\bar{\imath}(\mu)}{[\underline{\theta} - v, \underline{\theta} + v]}, ..., \overset{\bar{\jmath}(\mu)}{\left[\overline{\theta} - \nu, \overline{\theta} + \nu\right]}, ..., [., 1]$$

Now let us consider an alternative allocation $\mathbf{i}'(\theta)$ that generates a distribution identical to that generated by $\mathbf{i}^s(\theta)$, except the purchasing behavior of buyers in $[\underline{\theta} - v, \underline{\theta} + v]$ and $\left[\overline{\theta} - \nu, \overline{\theta} + \nu\right]$ is reversed, i.e.

$$..., [0, .], ..., \overset{\bar{\jmath}(\mu)}{[\underline{\theta} - v, \underline{\theta} + v]}, ..., \overset{\bar{\imath}(\mu)}{\left[\overline{\theta} - \nu, \overline{\theta} + \nu\right]}, ..., [., 1]$$

The difference between the total surplus under $\mathbf{i}^s(\theta)$ and $\mathbf{i}'(\theta)$ can be derived as

$$W(\mathbf{i}^s) - W(\mathbf{i}') = \left(\left(\alpha_{\bar{\imath}(\mu)} + \bar{\alpha}_{\bar{\imath}(\mu)}\left(\mu_{\bar{\jmath}(\mu)}\right)\right) - \left(\alpha_{\bar{\jmath}(\mu)} + \bar{\alpha}_{\bar{\jmath}(\mu)}\left(\mu_{\bar{\imath}(\mu)}\right)\right)\right) 8v\left(\underline{\theta} - \overline{\theta}\right) \quad \text{(A4)}$$

The first term is positive by definition of $\bar{\imath}(\boldsymbol{\mu})$ and the second term is negative since $\theta_1 < \theta_2$, thus $W(\mathbf{i}^s) - W(\mathbf{i}') < 0$. But this is a contradiction to $\mathbf{i}^s$ being optimal, so the distribution has to be of the form given in (A3), where $-1 \leqslant \theta_1 \leqslant 0 \leqslant \theta_2 \leqslant 1$.

(2) Recall from Definition 1 that if $\theta_1 < 0$, and $\theta_2 > 0$, the distribution in (A3) shuffled and instead if $\theta_1 = 0$ and/or $\theta_2 = 0$, the distribution is sorted. Next we will identify the conditions for which the former and the latter holds. We will start doing this by first solving the maximization problem subject to the resulting distribution being in $H^+ \cap U_h$ where $U_i = \left\{(\mu_l, \mu_h) \in U \mid \mu_i \geqslant \mu_j\right\}$ for $i \in \{l, h\}$. Thus let the total surplus maximizing solution conditional on $H^+ \cap U_h$ be $\mathbf{i}^s_{H^+}(\theta)$. So far we established that the distribution implied by $\mathbf{i}^s_{H^+}(\theta)$ has to be of the form in (A3). Notice that for $\boldsymbol{\mu} \in H^+ \cap U_h$, $\bar{\imath}(\boldsymbol{\mu}) = h$ and $\underline{i}(\boldsymbol{\mu}) = l$. Suppose $\mathbf{i}^s_{H^+}(\theta)$ actually implies that $\theta_1 < 0 < \theta_2$. Notice that since in $H^+ \cap U_h$

$\mu_h \geqslant \mu_l$ and by full market share $\mu_h + \mu_l = 1$, it has to be the case that $-\theta_1 > \theta_2$. Let us focus on the following two segments of the distribution

$$......, \overset{h}{[-\theta_2, 0]}, \overset{l}{[0, \theta_2]}, ...$$

Now consider an alternative allocation $\mathbf{i}'_{H+}(\theta)$ that generates a distribution identical to that generated by $\mathbf{i}^s_{H+}(\theta)$, except the purchasing behavior of buyers in $[-\theta_2, 0]$ and $[0, \theta_2]$ is reversed. Notice that with such a change the total measures of buyers buying h and l as implied by $\mathbf{i}^s_{H+}(\theta)$, (μ_h, μ_l) is unaltered. Now the difference in social surplus can be derived as

$$W(\mathbf{i}^s_{H+}) - W(\mathbf{i}'_{H+}) = -4\left((\alpha_h + \bar{\alpha}_h(\mu_l)) - (\alpha_l + \bar{\alpha}_l(\mu_h))\right)\theta_2^2 \tag{A5}$$

The first term $((\alpha_h + \bar{\alpha}_h(\mu_l)) - (\alpha_l + \bar{\alpha}_l(\mu_h))) > 0$, since $(\mu_h, \mu_l) \in H^+ \cap U_h$ and thus $W(\mathbf{i}^s_{H+}) - W(\mathbf{i}'_{H+}) < 0$. But this contradicts $\mathbf{i}^s_{H+}(\theta)$ being optimal, thus it has to be that either $\theta_1 = 0$ and/or $\theta_2 = 0$. This combined with the fact that $-\theta_1 > \theta_2$, implies either $\theta_2 = 0$ or $\theta_1 = \theta_2 = 0$. We can do the same analysis for region $L^- \cap U_l$, which would yield that $\mathbf{i}^s_{L^-}(\theta)$ has to satisfy either $\theta_1 = 0$ or $\theta_1 = \theta_2 = 0$. But notice that both of these imply that both $\mathbf{i}^s_{H+}(\theta)$ and $\mathbf{i}^s_{L^-}(\theta)$ has to be of the sorted form given below

$$\overset{l}{\left[-1, \widetilde{\theta}\right]}, \overset{h}{\left[\widetilde{\theta}, 1\right]} \tag{A6}$$

Let us find the optimal sorted-from allocation, which is equivalent to finding the optimal $\widetilde{\theta}$. Notice that under this rule $\mu_h = \frac{1-\widetilde{\theta}}{2}$ and $\mu_l = \frac{1+\widetilde{\theta}}{2}$. First let us restrict attention to

$\widetilde{\theta} \leqslant 0$, we can write the total surplus as a function of the unique switch point $\widetilde{\theta}$ as

$$W\left(\widetilde{\theta} \leqslant 0\right) = 4v - \frac{16}{3} + \int_{-1}^{\widetilde{\theta}} 4\theta\left(\alpha_l + \underline{\alpha}_l\left(\frac{1-\widetilde{\theta}}{2}\right)\right)d\theta \tag{A7}$$
$$+ \int_{\widetilde{\theta}}^{0} 4\theta\left(\alpha_h + \underline{\alpha}_h\left(\frac{1+\widetilde{\theta}}{2}\right)\right)d\theta + \int_{0}^{1} 4\theta\left(\alpha_h + \bar{\alpha}_h\left(\frac{1+\widetilde{\theta}}{2}\right)\right)d\theta$$

It is immediate that

$$\frac{dW\left(\widetilde{\theta} \leqslant 0\right)}{d\widetilde{\theta}} = 4\widetilde{\theta}\left(\left(\alpha_l + \underline{\alpha}_l\left(\frac{1-\widetilde{\theta}}{2}\right)\right) - \left(\alpha_h + \underline{\alpha}_h\left(\frac{1+\widetilde{\theta}}{2}\right)\right)\right) + \beta(\gamma,\varepsilon)\widetilde{\theta}^2 > 0 \tag{A8}$$

for all $\widetilde{\theta} < 0$, because $\beta(\gamma,\varepsilon) > 0$ and $\left(\alpha_l + \underline{\alpha}_l\left(\frac{1-\widetilde{\theta}}{2}\right)\right) - \left(\alpha_h + \underline{\alpha}_h\left(\frac{1+\widetilde{\theta}}{2}\right)\right) < 0$ due to $\widetilde{\theta} < 0$. We can do the same thing subject to $\widetilde{\theta} \geqslant 0$, which would yield $\frac{dW(\widetilde{\theta}\leqslant 0)}{d\widetilde{\theta}} < 0$. This implies the optimal switch point is $\widetilde{\theta}^s = 0$. Now we showed that the optimal sorted form allocation is balanced with $(\mu_l, \mu_h) = \left(\frac{1}{2}, \frac{1}{2}\right)$. This implies that the optimal distribution is either shuffled or balanced and sorted. Recall that we have shown in the region $H^+ \cap L^-$ the optimal allocation is sorted. Since$\left(\frac{1}{2}, \frac{1}{2}\right) \in H^+ \cap L^-$, this means the optimal allocation conditional on $H^+ \cap L^-$ is balanced and sorted.

(3) Finally we will look at the maximization problem conditional on regions L^+ and H^-. In these regions, as we described earlier in equation (A3) the optimal allocation implies the distribution is potentially shuffled. First notice that if $\rho < 3$, $L^+ = H^- = \emptyset$. Therefore when $\rho < 3$, the optimal distribution is balanced and sorted as in $H^+ \cap L^-$. Let us consider the case $\rho \geqslant 3$ and let us look at the total surplus maximizing allocations $\mathbf{i}^s_{L^+}(\theta)$ and $\mathbf{i}^s_{H^-}(\theta)$. We start with region L^+. We already showed that $\mathbf{i}^s_{L^+}(\theta)$ must be of the form given in (A3), where $-1 \leqslant \theta_1 \leqslant 0 \leqslant \theta_2 \leqslant 1$. We also know that $\bar{i}(\boldsymbol{\mu}) = \underline{i}(\boldsymbol{\mu}) = l$ since $\boldsymbol{\mu} \in$

L^+. Notice that this form of allocation implies $\mu_l = \frac{\theta_1}{2} + 1 - \frac{\theta_2}{2}$ and $\mu_h = \frac{\theta_2 - \theta_1}{2}$. We can write the total surplus as a function of θ_1 and μ_h using $\theta_2 = \theta_1 + 2\mu_h$:

$$\begin{aligned} W_{L^+}(\theta_1, \mu_h) &= 4v - \frac{16}{3} + 2\left(\theta_1^2 - 1\right)\left(-\varepsilon - \frac{1}{2}\mu_h(\gamma - \varepsilon)\right) \\ &\quad -2\theta_1^2\left(-\frac{1}{2}(1-\mu_h)(\gamma-\varepsilon)\right) + 2(\theta_1 + 2\mu_h)^2\left(\frac{1}{2}(1-\mu_h)(\gamma-\varepsilon) + \varepsilon\right) \\ &\quad +2\left(1 - (\theta_1 + 2\mu_h)^2\right)\left(\frac{1}{2}\mu_h(\gamma - \varepsilon)\right) \end{aligned} \tag{A9}$$

Notice that since there is full market share and we restrict the analysis to H^+, this implies $\mu_h \geqslant \frac{1}{2} + \frac{\varepsilon}{\gamma - \varepsilon}$. Also due to the two switch point structure described, $\theta_1 \leqslant 0$ and $\theta_1 + 2\mu_h \leqslant 1$. Now let us look at how the surplus changes in θ_1 for a given μ_h

$$\begin{aligned} \frac{\partial W_{L^+}(\theta_1, \mu_h)}{\partial \theta_1} &= -4\theta_1(2\mu_h - 1)(\gamma - \varepsilon) + 8\mu_h\left(-\frac{1}{2}(2\mu_h - 1)(\gamma - \varepsilon) + \varepsilon\right) \\ &\geqslant 8\left((\gamma - \varepsilon)\mu_h^2 + \left(-\frac{3}{2}(\gamma - \varepsilon) + \varepsilon\right)\mu_h + \frac{1}{2}(\gamma - \varepsilon)\right) \end{aligned} \tag{A10}$$

where the inequality is due to $\theta_1 + 2\mu_h \leqslant 1$ and rearranging. Careful analysis shows that $\frac{\partial W_{L^+}(\theta_1, \mu_h)}{\partial \theta_1} > 0$ for all θ_1 and μ_h when $7 - 4\sqrt{2} < \rho < 7 + 4\sqrt{2}$. If this is the case, for a given μ_h, the optimal choice of θ_1 should take the highest value possible, i.e. $1 - 2\mu_h$. But this reduces the problem to a problem of one switch point, i.e. sorted form, and we already showed that the optimal solution for that is for the switch point to be at zero. Thus for $\rho \leqslant 7 + 4\sqrt{2}$, the efficient distribution is sorted and balanced.

Now, we have to analyze the case $\rho > 7 + 4\sqrt{2}$, when the polynomial in (A10) has real roots given by $\mu_{h1} = \frac{3\rho - 5 - \sqrt{\rho^2 - 14\rho + 17}}{4(\rho - 1)}$, $\mu_{h2} = \frac{3\rho - 5 + \sqrt{\rho^2 - 14\rho + 17}}{4(\rho - 1)}$. Notice that when $\rho > 7 + 4\sqrt{2}$; for $\mu_h < \mu_{h1}$ or $\mu_h > \mu_{h2}$, $\frac{\partial W_{L^+}(\theta_1, \mu_h)}{\partial \theta_1} > 0$ for all θ_1 and thus given such μ_h, the optimal θ_1 is $1 - 2\mu_h$. But this means again the solution is of sorted form and as we showed the optimal sorted form distribution is balanced.

Now we know that the corner solutions in L^+ and H^- are suboptimal to the balanced and sorted distribution. We still need to look for interior solutions of shuffled form with $\theta_1 < 1-2\mu_h$, which might prove to perform better than the balanced and sorted distribution. Suppose $\mu_{h1} \leqslant \mu_h \leqslant \mu_{h2}$. Then since $\frac{\partial^2 W_{L^+}(\theta_1,\mu_h)}{\partial \theta_1^2} < 0$, we know that for a given μ_h the optimal θ_1 is the solution to $\frac{\partial W_{L^+}(\theta_1,\mu_h)}{\partial \theta_1} = 0$, which is

$$\theta_1^s(\mu_h) = -\mu_h + \frac{2\mu_h}{(2\mu_h - 1)(\rho - 1)} \tag{A11}$$

To find the interior solution let us look at how the total surplus changes in μ_h, when evaluated at $\theta_1(\mu_h)$.

$$\frac{\partial W_{L^+}(\theta_1(\mu_h), \mu_h)}{\partial \mu_h} = \frac{4\left((\gamma-\varepsilon)^2(2\mu_h-1)^2\left(-3\mu_h^2+\mu_h+\frac{1}{2}\right) - 4\varepsilon^2\mu_h(1-\mu_h)\right)}{(2\mu_h-1)^2(\gamma-\varepsilon)} \tag{A12}$$

Let $(\theta_1(\mu_h^s), \mu_h^s)$ be an interior solution to the maximization of $W_{L^+}(\theta_1, \mu_h)$. By the necessary conditions for a local maxima, $\mu_h^s \in S$ where

$$S = \left\{ \mu_h > \frac{1}{2} \mid \frac{dW_{L^+}(\theta_1^s(\mu_h), \mu_h)}{d\mu_h} = 0 \text{ and } \mu_h > \frac{1}{4} + \frac{\sqrt{15(\rho-1)^2+24}}{12(\rho-1)} \right\}$$

The first condition is the first order condition and the second condition is the local concavity condition that describes the region where the second order derivative is negative. It can be shown that the second order condition implies that either $S = \emptyset$ or $|S| = 1$. Let us first find then when $|S| = 1$, i.e. when such a local maximum exists. First notice from equation

(A12) that if $\frac{dW_{L^+}(\theta_1(\mu_h^s),\mu_h^s)}{d\mu_h} = 0$, $\left(-3\mu_h^{s2} + \mu_h^s + \frac{1}{2}\right) > 0$,but this implies

$$\frac{\partial^2\left(\frac{(\gamma-\varepsilon)(2\mu_h-1)^2}{4\varepsilon^2}W_{L^+}\left(\theta_1\left(\mu_h^s\right),\mu_h^s\right)\right)}{\partial\mu_h\partial\rho} = 2\left(\rho-1\right)\left(2\mu_h-1\right)^2\left(-3\mu_h^2+\mu_h+\frac{1}{2}\right) > 0 \quad \text{(A13)}$$

Therefore using implicit function theorem with (A13) and the second order optimality condition

$$\frac{d\mu_h^s}{d\rho} > 0 \quad \text{(A14)}$$

Then, since $\frac{1}{4} + \frac{\sqrt{15(\rho-1)^2+24}}{12(\rho-1)}$ is decreasing in ρ, finding when $|S| = 1$ is equivalent to finding the $\widehat{\rho} > 7 + 4\sqrt{2}$ that solves

$$\left.\frac{dW_{L^+}\left(\theta_1\left(\mu_h\right),\mu_h\right)}{d\mu_h}\right|_{\mu_h=\frac{1}{4}+\frac{\sqrt{15(\widehat{\rho}-1)^2+24}}{12(\widehat{\rho}-1)}} = 0 \quad \text{(A15)}$$

Simple algebra shows that $\widehat{\rho} \sim 23.84$. Thus for $\rho \leqslant \widehat{\rho}$, no shuffled form solution exists, which implies that the optimal allocation is the balanced and sorted one.

Now, we know that an interior shuffled solution given by μ_h^s exists for $\rho > \widehat{\rho}$. A simple algebraic check using the conditions on μ_h^s also shows that $\mu_{h1} \leq \mu_h^s \leq \mu_{h2}$. We have to compare it to the sorted and balanced solution and see for $\rho > \widehat{\rho}$ when which one dominates. Let us look at the difference between the total surplus achieved under the optimal shuffled form $\left(\theta_1^s\left(\mu_h^s\right),\mu_h^s\right)$ and the balanced and sorted form, normalized by ε. First simple algebra shows that $\left.\frac{W_{L^+}(\theta_1^s(\mu_h^s),\mu_h^s)-W(\widetilde{\theta}=0)}{\varepsilon}\right|_{\rho=\widehat{\rho}} < 0$.and $\lim_{\rho\to\infty}\frac{W_{L^+}(\theta_1^s(\mu_h^s),\mu_h^s)-W(\widetilde{\theta}=0)}{\varepsilon} > 0$. Then we can look at the first derivative of this difference with respect to ρ, using envelope theorem

$$\frac{d\left(\frac{W_{L^+}(\theta_1^s(\mu_h^s),\mu_h^s)-W(\tilde{\theta}=0)}{\varepsilon}\right)}{d\rho} = -\frac{1}{8}\left(2\mu_h^s-1\right)\left(2\mu_h^{s2}-1\right) - \frac{\mu_h^{s2}}{(\rho-1)^2\left(2\mu_h^s-1\right)} \tag{A16}$$

Using the conditions on $\widehat{\rho}$,optimality of μ_h^s and the fact that $\frac{d\mu_h^s}{d\rho} > 0$, it can be shown that $\left.\frac{d\left(\frac{W_{L^+}(\theta_1(\mu_h^s),\mu_h^s)-W(\tilde{\theta}=0)}{\varepsilon}\right)}{d\rho}\right|_{\widehat{\rho}} > 0$, $\frac{d^2\left(\frac{W_{L^+}(\theta_1(\mu_h^s),\mu_h^s)-W(\tilde{\theta}=0)}{\varepsilon}\right)}{d\rho^2} > 0$. But this clearly means that $\frac{d\left(\frac{W_{L^+}(\theta_1^s(\mu_h^s),\mu_h^s)-W(\tilde{\theta}=0)}{\varepsilon}\right)}{d\rho} > 0$ for all $\rho > \widehat{\rho}$. These together with the initial condition and the limit condition above imply that there is a unique $\rho > \widehat{\rho}$ that satisfies $\frac{W_{L^+}(\theta_1^s(\mu_h^s),\mu_h^s)-W(\tilde{\theta}=0)}{\varepsilon} = 0$. Call that ρ_s.

Notice that if we try to find the interior solution in H^-, we will get a symmetric interior solution where the critical ρ is still ρ_s. ∎

Proof of Proposition 1.2. The previous proof shows that for $\rho \geqslant \widehat{\rho}$, one of the efficient allocations is $(\mu_h^s, 1-\mu_h^s) \in L^+$. But this means $\mu_h^s - (1-\mu_h^s) \geqslant \frac{2}{\rho-1} >> 0$ for all $\rho << \infty$. Hence at $\rho = \widehat{\rho} << \infty$, the level of unbalance is $\frac{\mu_h^s}{(1-\mu_h^s)} >> 1$, which means that there is a discontinuous jump at $\rho = \widehat{\rho}$. The same argument holds for the symmetric allocation.

Remember that ρ_s identifies the point at which the efficient solution becomes an interior point in L^+ and H^-. As we showed in the proof of the previous proposition in equation (A14), for the interior shuffled form solution in L^+, $\frac{d\mu_h^s}{d\rho} > 0$. Since $\mu_l^* = 1-\mu_h^s$, this implies $\frac{\mu_h^s}{\mu_l^s}$, i.e. the degree of unbalance, is increasing in ρ.

To show the change in the shuffling structure, let us look at θ_1^* and $\theta_1^* + 2\mu_h^s$, which are the two switch points on each side of 0. Using (A11),

$$\frac{d\theta_1^s(\mu_h^s)}{d\rho} = -\frac{2\mu_h^s}{(2\mu_h^s-1)(\rho-1)^2} + \left(-1 - \frac{1}{(\rho-1)}\frac{2}{(2\mu_h^s-1)^2}\right)\frac{d\mu_h^s}{d\rho} < 0 \tag{A17}$$

where the inequality is due to $\frac{d\mu_h^s}{d\rho} > 0$.

Using the implicit function theorem definition of $\frac{d\mu_h^s}{d\rho}$ and the first order condition, it can be derived that

$$\begin{aligned}\frac{d\left(\theta_1^s(\mu_h^s)+2\mu_h^s\right)}{d\rho} &= \frac{2\mu_h^s\left(-1+\frac{4(1-\mu_h^s)\left((2\mu_h^s-1)^2(\rho-1)-2\right)}{\left((\rho-1)^2(24\mu_h^{s2}-12\mu_h^s-1)-4\right)(2\mu_h^s-1)^2}\right)}{(2\mu_h^s-1)(\rho-1)^2} \\ &< \frac{2\mu_h^s}{(2\mu_h^s-1)(\rho-1)^2}\left(-1+\frac{4(\rho-1)(1-\mu_h^s)}{(\rho-1)^2(24\mu_h^{s2}-12\mu_h^s-1)-4}\right)\end{aligned} \tag{A18}$$

Using the second order condition and the fact that μ_h^s is increasing, it can be shown that $-1+\frac{4(\rho-1)(1-\mu_h^s)}{(\rho-1)^2(24\mu_h^{s2}-12\mu_h^s-1)-4} < 0$ for all $\rho > \widetilde{\rho} \sim 47.48$ where $\widetilde{\rho}$ is the solution to $(2\mu_h^s(\widetilde{\rho})-1)^2(\widetilde{\rho}-1)-2=0$. Hence $\frac{d(\theta_1^s+2\mu_h^s)}{d\rho} < 0$ for all $\rho \geqslant \rho_s$.

Since $\frac{d(\theta_1^s+2\mu_h^s)}{d\rho} < 0$ and $\frac{d\theta_1^s}{d\rho} < 0$, it is clear that $\frac{d\left(\frac{1-(\theta_1^s+2\mu_h^s)}{\theta_1^s+1}\right)}{d\rho} > 0$, which proves that the degree of shuffling is increasing in ρ.

The limit case where $\rho \to \infty$ can easily be derived from (A11) and (A12) as

$$\lim_{\rho\to\infty}\theta_1^s(\mu_h) = -\mu_h$$

since we already showed that $\mu_h > \frac{1}{2}$

$$\lim_{\varepsilon\to 0}\frac{dW_{L^+}(\theta_1^s(\mu_h),\mu_h)}{d\mu_h} = \frac{4}{(2\mu_h-1)^2\gamma}\gamma^2(2\mu_h-1)^2\left(-3\mu_h^2+\mu_h+\frac{1}{2}\right)$$

Optimality conditions give $\frac{1}{6}+\frac{1}{6}\sqrt{7}$ as the maximizer. ∎

Proof of Propositions 1.3-1.6, [Case 1.] We divide the ρ space into four and find the maximizing (μ_h, μ_l) for each subspace.

[CASE 1: $\rho<3$] First notice that for $\rho<3$, $L^{+}=LF^{+}=H^{-}=HF^{-}=\emptyset$. In other words $H^{+}=FL^{+}=L^{-}=FH^{-}=U$. Let us first look at the $\boldsymbol{\mu}=(\mu_l,\mu_h)\in U$ such that $p_l(\boldsymbol{\mu}), p_h(\boldsymbol{\mu})\geqslant c$. Due to the single crossing property this results in the following segmentation, which we will call "Segmentation 1"

$$[-1, \overset{l}{\theta_1}(\boldsymbol{\mu})], [\theta_1\overset{F,l}{(\boldsymbol{\mu})}, 0], [0, \overset{F,h}{\theta_2}(\boldsymbol{\mu})], [\theta_2\overset{h}{(\boldsymbol{\mu})}, 1] \tag{A19}$$

where θ_1 and θ_2 are given by the following incentive constraints, which are satisfied by equality.

$$\begin{aligned} 4\theta_1(\boldsymbol{\mu})(-\frac{1}{2}\mu_h\beta(\gamma,\varepsilon)) &= p_l(\boldsymbol{\mu})-c \\ 4\theta_2(\boldsymbol{\mu})(\frac{1}{2}\mu_l\beta(\gamma,\varepsilon)) &= p_h(\boldsymbol{\mu})-c \end{aligned} \tag{A20}$$

Notice that this distribution implies $\mu_l=\frac{\theta_1(\boldsymbol{\mu})+1}{2}$ and $\mu_h=\frac{1-\theta_2(\boldsymbol{\mu})}{2}$. Substituting the market shares for marginal buyers we can write the reduced profits for Segmentation 1 as

$$\pi_M(\mu_h,\mu_l)=4\beta(\gamma,\varepsilon)\,\mu_l\mu_h(1-\mu_l-\mu_h) \tag{A21}$$

The first order derivatives with respect to μ_h and μ_l imply that the optimal measures are

$$\mu_h = \frac{1-\mu_l}{2},\ \mu_l = \frac{1-\mu_h}{2} \tag{A22}$$

since $\beta(\gamma, \varepsilon) > 0$. Combining these two gives us

$$\arg \max_{(\mu_h,\mu_l)\in U, p_l(\mu), p_h(\mu) \geqslant c} \pi_M(\mu_h, \mu_l) = \left(\frac{1}{3}, \frac{1}{3}\right) \tag{A23}$$

with a profit level of $\pi_M\left(\frac{1}{3}, \frac{1}{3}\right) = \frac{4}{27}\beta(\gamma, \varepsilon)$ when we restrict attention to positive prices .

Now let us look into what happens if the seller actually charges prices below marginal cost. We can immediately eliminate the case where the price of both products are below c, because that would bring negative profits. But suppose $\boldsymbol{\mu}$ is such that $p_h(\boldsymbol{\mu}) < c$ and $p_l(\boldsymbol{\mu}) > c$. This results in the following segmentation, which we will call "Segmentation 2"

$$\overset{l}{[-1, \theta_1(\boldsymbol{\mu})]}, \overset{F,l}{[\theta_1(\boldsymbol{\mu}), \theta_2(\boldsymbol{\mu})]}, \overset{h}{[\theta_2(\boldsymbol{\mu}), 0]}, \overset{h}{[0, 1]} \tag{A24}$$

θ_1 and θ_2 are given by the following indifference conditions:

$$\begin{aligned} 4\theta_1(\boldsymbol{\mu})(-\frac{1}{2}\mu_h\beta(\gamma,\varepsilon)) &= p_l(\boldsymbol{\mu}) - c \\ 4\theta_2(\boldsymbol{\mu})(\varepsilon - \frac{1}{2}\mu_l\beta(\gamma,\varepsilon)) &= p_h(\boldsymbol{\mu}) - c \end{aligned} \tag{A25}$$

Notice that this segmentation implies $\mu_l = \frac{\theta_1(\boldsymbol{\mu})+1}{2}$ and $\mu_h = \frac{1-\theta_2(\boldsymbol{\mu})}{2} > \frac{1}{2}$. Substituting the market shares for marginal buyers we can write the reduced profits for Segmentation 2 as

$$\pi_M(\mu_h,\mu_l) = 4(1-2\mu_l)(\frac{1}{2}\mu_h\beta(\gamma,\varepsilon))\mu_l + 4(1-2\mu_h)(\varepsilon - \frac{1}{2}\mu_l\beta(\gamma,\varepsilon))\mu_h \tag{A26}$$

The second order derivative is negative and the first order condition with respect to μ_l implies that given μ_h the optimal μ_l is

$$\mu_l = \frac{\mu_h}{2} \tag{A27}$$

Let us plug this back in the profit function and look at how it changes with respect to μ_h.

$$\pi_M\left(\mu_h,\frac{\mu_h}{2}\right) = 4(1-\mu_h)(\frac{1}{2}\mu_h\beta(\gamma,\varepsilon))\frac{\mu_h}{2} + 4(1-2\mu_h)(\varepsilon - \frac{1}{2}\frac{\mu_h}{2}\beta(\gamma,\varepsilon))\mu_h \tag{A28}$$

$$\frac{\partial\pi_M(\mu_h,\mu_l(\mu_h))}{\partial\mu_h} = 3\beta(\gamma,\varepsilon)\mu_h^2 - 16\mu_h\varepsilon + 4\varepsilon \tag{A29}$$

The second derivative $\frac{\partial^2\pi_M(\mu_h,\mu_l(\mu_h))}{\partial\mu_h^2} < 0$ and $\left.\frac{\partial\pi_M(\mu_h,\mu_l(\mu_h))}{\partial\mu_h}\right|_{\mu_h=\frac{1}{2}} < 0$ for $\rho \leqslant 3$, since $\beta(\gamma,\varepsilon) < 2\varepsilon$ if $\rho \leqslant 3$. But then $\frac{\partial\pi_M(\mu_h,\mu_l(\mu_h))}{\partial\mu_h} < 0$ for all $\mu_h \geqslant \frac{1}{2}$. Hence

$$\arg\max_{(\mu_h,\mu_l)\in U, p_l(\mu)>c, p_h(\mu)<c} \pi_M(\mu_h,\mu_l) = \left(\frac{1}{2},\frac{1}{4}\right) \text{ for } \rho \leqslant 3 \tag{A30}$$

which generates a profit of $\pi_M\left(\frac{1}{2},\frac{1}{4}\right) = \frac{1}{8}\beta(\gamma,\varepsilon)$. Now if we look at the difference between this segmentation and the previous one we get $\pi_M\left(\frac{1}{3},\frac{1}{3}\right) - \pi_M\left(\frac{1}{2},\frac{1}{4}\right) > 0$. This proves that

$$\left(\mu_h^{*1},\mu_l^{*1}\right) = \arg\max_{(\mu_h,\mu_l)\in U} \pi_M(\mu_h,\mu_l) = \left(\frac{1}{3},\frac{1}{3}\right) \text{ for all } \rho \leqslant 3 \tag{A31}$$

with $p_l\left(\mu_h^{*1},\mu_l^{*1}\right), p_h\left(\mu_h^{*1},\mu_l^{*1}\right) \geqslant c$. Also notice that we can think of $\left(\mu_h^{*1},\mu_l^{*1}\right)$ as the maximizer of Segmentation 1, which we will use later.

∎

Proof of Propositions 1.3-1.6, [Case 2.] [CASE 2: $3 < \rho \leqslant 5$]. From this point on we will refer to the following graph:

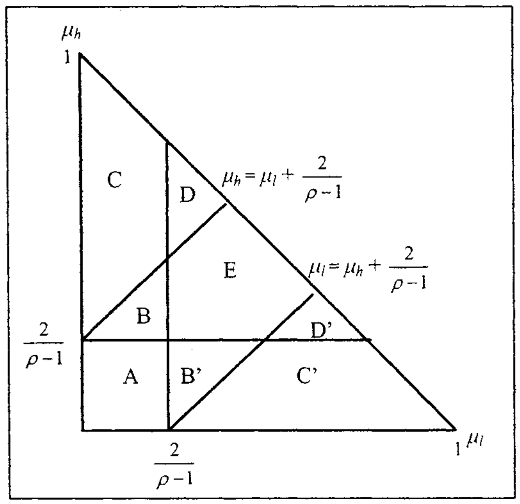

Figure 1.12: Ranking Division

In this graph $A = H^+ \cap FL^+ \cap L^- \cap FH^-$, $B = H^+ \cap LF^+ \cap L^- \cap FH^-$, $C = L^+ \cap LF^+ \cap L^- \cap FH^-$, $D = L^+ \cap LF^+ \cap L^- \cap HF^-$, $E = H^+ \cap LF^+ \cap L^- \cap HF^-$ and B', C', D' are defined symmetrically.

The first thing to notice is that for $\rho \leqslant 5$, $D = D' = E = \emptyset$. Region A is what we covered in the previous case. Thus we have already found the optimal solution there, which was $\left(\mu_h^{*1},\mu_l^{*1}\right) = \left(\frac{1}{3},\frac{1}{3}\right) \in A$.

The next thing to notice is that for $\rho \leqslant 5$, $\left\{(\mu_l,\mu_h) \in B \cup C \mid \mu_h < \frac{1}{2}\right\} = \emptyset$. Thus we restrict attention to $\mu_h \geqslant \frac{1}{2}$. Let us start with region B. The first thing to notice is that since $B \subset FH^-$, to achieve $\mu_h \geqslant \frac{1}{2}$ it has to be the case that $p_h(\boldsymbol{\mu}) \leqslant c$. Thus we will be looking at prices $p_l(\boldsymbol{\mu}) \geqslant c$, $p_h(\boldsymbol{\mu}) \leqslant c$. This pricing scheme generates Segmentation 2 as given by (A24). We already found the reduced form profit for this in CASE 1, which was given in (A26). We can substitute $\beta(\gamma, \varepsilon) = \gamma - \varepsilon$ in (A26), since $\rho \geqslant 3$. Then the first order derivative with respect to μ_l achieves zero at $\mu_l = \frac{\mu_h}{2}$. Since the second order derivative is negative,we can say that

$$\arg \max_{\mu_l \text{ such that } (\mu_l, \mu_h) \in B} \pi_M(\mu_h, \mu_l) = \min\left\{\frac{\mu_h}{2}, 1 - \mu_h\right\} = \begin{cases} 1 - \mu_h \text{ if } \mu_h \geqslant \frac{2}{3} \\ \frac{\mu_h}{2} \text{ otherwise.} \end{cases} \tag{A32}$$

Let us also find the maximizing μ_l for each μ_h in region C and then find the maximizing μ_h in region $B \cup C$. Notice that also in region C, $\mu_h \geqslant \frac{2}{\rho - 1} \geqslant \frac{1}{2}$ since $\rho \leqslant 5$. By the same reasoning as in region B, $p_h(\boldsymbol{\mu}) \leqslant c$. In this region prices $p_l(\boldsymbol{\mu}) \geqslant c$, $p_h(\boldsymbol{\mu}) \leqslant c$ generates the following segmentation, which we will call "Segmentation 3",

$$\overset{l}{[-1, \theta_1(\boldsymbol{\mu})]}, \overset{F,l}{[\theta_1(\boldsymbol{\mu}), \theta_2(\boldsymbol{\mu})]}, \overset{h}{[\theta_2(\boldsymbol{\mu}), 0]}, \overset{h}{[0, \theta_3(\boldsymbol{\mu})]}, \overset{l}{[\theta_3(\boldsymbol{\mu}), 1]} \tag{A33}$$

where $-1 \leqslant \theta_1(\boldsymbol{\mu}) \leqslant \theta_2(\boldsymbol{\mu})$ and $0 \leqslant \theta_3(\boldsymbol{\mu}) \leqslant 1$ are given by the following incentive constraints:

$$\begin{aligned} 4\theta_1(\boldsymbol{\mu})\left(-\frac{1}{2}\mu_h(\gamma - \varepsilon)\right) &= p_l(\boldsymbol{\mu}) - c \\ 4\theta_2(\boldsymbol{\mu})\left(\varepsilon - \frac{1}{2}\mu_l(\gamma - \varepsilon)\right) &= p_h(\boldsymbol{\mu}) - c \\ 4\theta_3(\boldsymbol{\mu})\left(\frac{1}{2}(\mu_h - \mu_l)(\gamma - \varepsilon) - \varepsilon\right) &\leqslant p_l(\boldsymbol{\mu}) - p_h(\boldsymbol{\mu}) \end{aligned} \tag{A34}$$

Note that if $\theta_3(\boldsymbol{\mu}) < 1$, the last inequality is satisfied with equality. First we will show that $\theta_3(\boldsymbol{\mu}) = 1$. Suppose $\theta_3(\boldsymbol{\mu}) < 1$, then this implies $\mu_h = \frac{\theta_3(\boldsymbol{\mu}) - \theta_2(\boldsymbol{\mu})}{2}$ and $\mu_l = \mathbf{1} + \frac{\theta_1(\boldsymbol{\mu}) - \theta_3(\boldsymbol{\mu})}{2}$. From these expressions we can write that $\theta_2(\boldsymbol{\mu}) = \theta_3(\boldsymbol{\mu}) - 2\mu_h$ and $\theta_1(\boldsymbol{\mu}) = 2(\mu_l - \mathbf{1}) + \theta_3(\boldsymbol{\mu})$. From the first two incentive constraints and substituting for $\theta_1(\boldsymbol{\mu})$ and $\theta_2(\boldsymbol{\mu})$ we can induce that $p_l(\boldsymbol{\mu}) - p_h(\boldsymbol{\mu}) = 4(2(\mu_l - \mathbf{1}) + \theta_3(\boldsymbol{\mu}))(-\frac{1}{2}\mu_h(\gamma - \varepsilon)) - 4(\theta_3(\boldsymbol{\mu}) - 2\mu_h)(\varepsilon - \frac{1}{2}\mu_l(\gamma - \varepsilon))$

Now let us look at the LHS-RHS of the third incentive constraint substituting for $p_l(\boldsymbol{\mu}) - p_h(\boldsymbol{\mu})$ from (A35),

$$\begin{aligned} & 4\theta_3(\boldsymbol{\mu})((\mu_h - \mu_l)(\gamma - \varepsilon)) + 4(\mu_l - \mathbf{1})\mu_h(\gamma - \varepsilon) - 8\mu_h(\varepsilon - \frac{1}{2}\mu_l(\gamma - \varepsilon)) \\ < \; & 4((\gamma - \varepsilon)\mu_l(2\mu_h - 1) - 2\mu_h\varepsilon) < 8\varepsilon(\mu_h - 1) < 0 \end{aligned}$$

The first inequality is due to $\theta_3(\boldsymbol{\mu}) < 1$ and the second to last is due to $\mu_l \leqslant \frac{2\varepsilon}{\gamma - \varepsilon}$. But this means, the third incentive constraint is not satisfied with equality, thus it has to be the case that $\theta_3(\boldsymbol{\mu}) = 1$. Therefore, Segmentation 3 in (A33) is reduced to Segmentation 2 in (A24), which is the same as the segmentation in region B. But we already found the maximizing μ_l for each μ_h in equation (A32).

Now we will maximize the profit with respect to μ_h given that μ_l is chosen optimally for each μ_h. Let us look at the profits from both possibilities as given by (A32) and first restrict attention to $\mu_h \geqslant \frac{2}{3}$, which implies for each μ_h the optimal solution is full market share and yields a profit

$$\pi_M(\mu_h, 1 - \mu_h) = 4\mu_h(1 - \mu_h)(\gamma - \varepsilon)(2\mu_h - 1) + 4\mu_h(1 - 2\mu_h)\varepsilon \tag{A36}$$

Let us find the maximum of this function

$$\frac{d\pi_M\left(\mu_h, 1-\mu_h\right)}{d\mu_h} = 4\left(-6\left(\gamma-\varepsilon\right)\mu_h^2 + 2\left(3\gamma-5\varepsilon\right)\mu_h + 2\varepsilon-\gamma\right) \tag{A37}$$

Let $\left(\mu_h^{*2f}, 1-\mu_h^{*2f}\right) \equiv \arg\max_{\mu_h \geqslant \frac{2}{3} \text{ and } \mu_h \in B\cup C} \pi_M\left(\mu_h, 1-\mu_h\right)$. Looking at the roots of this polynomial and the second order derivative reveals

$$\mu_h^{*2f} = \begin{cases} \frac{2}{3} \text{ if } \rho \leqslant 6 \\ \frac{1}{2} - \frac{2-\sqrt{3\rho^2-12\rho+13}}{6(\rho-1)} \text{ if } 6 < \rho \leqslant \frac{11}{2} + \frac{1}{2}\sqrt{17} \\ 1 - \frac{2}{\rho-1} \text{ otherwise} \end{cases} \tag{A38}$$

which we can also refer to as the Segmentation 2 full market share maximizer.

Since we are analyzing the case where $\rho \leqslant 5$, this means $\arg\max_{\mu_h \geqslant \frac{2}{3}} \pi_M\left(\mu_h, 1-\mu_h\right) = \frac{2}{3}$ giving a profit of $\pi_M\left(\frac{2}{3}, \frac{1}{3}\right) = \frac{8}{27}\left(\gamma-\varepsilon\right) - \frac{8}{9}\varepsilon$. Again comparison shows that for all $\rho \leqslant 6$, $\pi_M\left(\frac{1}{3}, \frac{1}{3}\right) \geqslant \pi_M\left(\frac{2}{3}, \frac{1}{3}\right)$ and hence $\left(\mu_h^{*1}, \mu_l^{*1}\right)$performs better than $\left(\mu_h^{*2f}, \mu_l^{*2f}\right)$.

Now let us look at the maximizing μ_h, for $\mu_h < \frac{2}{3}$. The relevant profit function is

$$\pi_M\left(\mu_h, \frac{\mu_h}{2}\right) = \mu_h^3\left(\gamma-\varepsilon\right) + 4\mu_h(1-2\mu_h)\varepsilon \tag{A39}$$

The first derivative of this function with respect to μ_h is

$$\frac{d\pi_M\left(\mu_h, \frac{\mu_h}{2}\right)}{d\mu_h} = 3\left(\gamma-\varepsilon\right)\mu_h^2 - 16\varepsilon\mu_h + 4\varepsilon \tag{A40}$$

Notice that when $\rho > \frac{19}{3}$ there are no real roots and the derivative is always increasing. For the purpose of CASE 2, suppose $\rho \leqslant \frac{19}{3}$ and let the roots be $\mu_{h1} < \mu_{h2}$. It can be shown that $\mu_{h1} \leqslant \frac{1}{2}$. Thus we will be looking for $\mu_h \geqslant \mu_{h1}$. Notice that $\pi_M\left(\mu_h, \frac{\mu_h}{2}\right)$ is quasi-convex for $\mu_h > \mu_{h1}$. Simple algebra reveals that $\mu_{h2} \lessgtr \frac{2}{3}$ as $\rho \gtrless 6$. But this means for $\rho \leqslant 5$, since $\mu_{h1} \leqslant \frac{1}{2} < \frac{2}{3} \leqslant \mu_{h2}$, $\max_{\mu_h \leqslant \frac{2}{3}} \pi_M\left(\mu_h, \frac{\mu_h}{2}\right) \leqslant \pi_M\left(\frac{1}{2}, \frac{1}{4}\right)$. We have already shown in CASE 1 that $\pi_M\left(\frac{1}{2}, \frac{1}{4}\right) < \pi_M\left(\frac{1}{3}, \frac{1}{3}\right)$, which implies that the maximizer of region A,i.e. $\left(\mu_h^{*1}, \mu_l^{*1}\right)$ performs better.

Therefore, for $\rho \leqslant 5$,

$$\left(\mu_h^{*1}, \mu_l^{*1}\right) = \arg \max_{(\mu_h, \mu_l) \in U} \pi_M\left(\mu_h, \mu_l\right) = \left(\frac{1}{3}, \frac{1}{3}\right) \tag{A41}$$

∎

Proof of Propositions 1.3-1.6, [Case 3.] [CASE 3: $5 < \rho \leqslant 7$] Notice that in this case, $D = D' = \emptyset$ and all other regions are non-empty. The analysis for region A is as before. For $\rho > 5$, $\left\{(\mu_h, \mu_l) \in B \cup C \mid \mu_h < \frac{1}{2}\right\} \neq \emptyset$. First we will restrict attention to $\left\{(\mu_h, \mu_l) \in B \mid \mu_h \geqslant \frac{1}{2}\right\} \cup \left\{(\mu_h, \mu_l) \in C \mid \mu_h \geqslant \frac{1}{2}\right\}$,which we will respectively refer to as $\overline{B} \cup \overline{C}$. (A38) and (A40) in the previous case imply that, for $\rho \leqslant 6$ the analysis is the same and thus the maximizer in regions $A \cup \overline{B} \cup \overline{C}$ is still $\left(\mu_h^{*1}, \mu_l^{*1}\right) = \left(\frac{1}{3}, \frac{1}{3}\right)$. However, when $\rho > 6$, as we showed in the Segmentation 2 full market share maximizer in (A38) $\arg\max_{\mu_h \geqslant \frac{2}{3}} \pi_M\left(\mu_h, 1-\mu_h\right) = \frac{1}{2} - \frac{2-\sqrt{3\rho^2-12\rho+13}}{6(\rho-1)}$. Simple algebra will show that for $\pi_M\left(\frac{1}{3}, \frac{1}{3}\right) \lesseqgtr \pi_M\left(\frac{1}{2} - \frac{2-\sqrt{3\rho^2-12\rho+13}}{6(\rho-1)}, \frac{1}{2} + \frac{2-\sqrt{3\rho^2-12\rho+13}}{6(\rho-1)}\right)$ as $\rho \lesseqgtr \widehat{\rho}_1 \sim 6.92$. Combining this with our findings for $\mu_h < \frac{2}{3}$ in CASE 2 implies that

$$\arg \max_{(\mu_h, \mu_l) \in A \cup \overline{B} \cup \overline{C}} \pi_M\left(\mu_h, \mu_l\right) = \begin{cases} \left(\frac{1}{3}, \frac{1}{3}\right) \text{ if } \rho < \widehat{\rho}_1 \\ \left(\frac{1}{2} - \frac{2-\sqrt{3\rho^2-12\rho+13}}{6(\rho-1)}, \frac{1}{2} + \frac{2-\sqrt{3\rho^2-12\rho+13}}{6(\rho-1)}\right) \text{ if } 7 \geqslant \rho \geqslant \widehat{\rho}_1 \end{cases} \tag{A42}$$

Now let us look at what happens in $\{(\mu_h, \mu_l) \in B \cup C \mid \mu_h < \frac{1}{2}\}$. Let us first focus on $\underline{C} = \{(\mu_h, \mu_l) \in C \mid \mu_h < \frac{1}{2}\}$. Notice that since for all $(\mu_h, \mu_l) \in \underline{C}$, $\frac{1}{2} > \mu_h \geqslant \mu_l$, it has to be the case that $c < p_h \leqslant p_l$. But this kind of pricing scheme results in Segmentation 1 given in (A19) again and we know that with this segmentation the optimal profit is achieved at $(\mu_h^{*1}, \mu_l^{*1}) = (\frac{1}{3}, \frac{1}{3})$, which is already covered in region A.

Now let us look at $\underline{B} = \{(\mu_h, \mu_l) \in C \mid \mu_h < \frac{1}{2}\}$. By a similar reasoning, since for all $(\mu_h, \mu_l) \in \underline{C}$, $\frac{1}{2} > \mu_h \geqslant \mu_l$, it has to be the case that $p_h(\boldsymbol{\mu}) \leqslant p_l(\boldsymbol{\mu})$. Notice that it is not necessary that $c < p_h(\boldsymbol{\mu})$. Suppose $c \geqslant p_h(\boldsymbol{\mu})$.This results in Segmentation 3 as we described in (A33) and using the same reasoning as before we can show that $\theta_3(\boldsymbol{\mu}) = 1$. This implies $\mu_h = \frac{\theta_3(\mu)-\theta_2(\mu)}{2} \geqslant \frac{1}{2}$, since $\theta_2(\boldsymbol{\mu}) \leqslant 0$. But this contradicts $\mu_h < \frac{1}{2}$. Therefore, to generate a $(\mu_h, \mu_l) \in \underline{B}$, it has to be the case that $c < p_h(\boldsymbol{\mu})$. Now $c < p_h(\boldsymbol{\mu}) \leqslant p_l(\boldsymbol{\mu})$, generates the following segmentation, which we call "Segmentation 4".

$$\overset{l}{[-1, \theta_1(\boldsymbol{\mu})]}, \overset{F,l}{[\theta_1(\boldsymbol{\mu}), 0]}, \overset{F,h}{[0, \theta_2(\boldsymbol{\mu})]}, \overset{h}{[\theta_2(\boldsymbol{\mu}), \theta_3(\boldsymbol{\mu})]}, \overset{l}{[\theta_3(\boldsymbol{\mu}), 1]} \tag{A43}$$

where $-1 \leqslant \theta_1(\boldsymbol{\mu}) \leqslant 0 \leqslant \theta_2(\boldsymbol{\mu}) \leqslant \theta_3(\boldsymbol{\mu}) \leqslant 1$ are given by the following incentive constraints:

$$\begin{aligned} 4\theta_1(\boldsymbol{\mu})(-\frac{1}{2}\mu_h(\gamma - \varepsilon)) &= p_l(\boldsymbol{\mu}) - c \\ 4\theta_2(\boldsymbol{\mu})(\frac{1}{2}\mu_l(\gamma - \varepsilon)) &= p_h(\boldsymbol{\mu}) - c \\ 4\theta_3(\boldsymbol{\mu})(\frac{1}{2}(\mu_h - \mu_l)(\gamma - \varepsilon) - \varepsilon) &\leqslant p_l(\boldsymbol{\mu}) - p_h \end{aligned} \tag{A44}$$

Again, if $\theta_3(\boldsymbol{\mu}) < 1$, the last inequality is satisfied with equality. Let us try to do the same analysis that we did before. Given this distribution, $\mu_h = \frac{\theta_3(\mu)-\theta_2(\mu)}{2}$ and $\mu_l = \mathbf{1} + \frac{\theta_1(\mu)-\theta_3(\mu)}{2}$. From these expressions we can write that $\theta_2(\boldsymbol{\mu}) = \theta_3(\boldsymbol{\mu}) - \mathbf{2}\mu_h$ and $\theta_1(\boldsymbol{\mu}) = 2(\mu_l - \mathbf{1}) + \theta_3(\boldsymbol{\mu})$.

From the first two incentive constraints in (A44) we can induce that

$$p_l(\boldsymbol{\mu}) - p_h(\boldsymbol{\mu}) = 4\left(2\left(\mu_l - \mathbf{1}\right) + \theta_3(\boldsymbol{\mu})\right)\left(-\frac{1}{2}\mu_h(\gamma - \varepsilon)\right) - 4\left(\theta_3(\boldsymbol{\mu}) - \mathbf{2}\mu_h\right)\left(\frac{1}{2}\mu_l(\gamma - \varepsilon)\right) \tag{A45}$$

Now suppose $\theta_3(\mathbf{p}) < 1$. Once we use this price difference, LHS-RHS of the third incentive constraint reduces to

$$4\theta_3(\boldsymbol{\mu})\left(\mu_h(\gamma - \varepsilon) - \varepsilon\right) - 4\mu_h(\gamma - \varepsilon) < -4\varepsilon$$

where, the inequality is due to $\theta_3(\boldsymbol{\mu}) < 1$. This means the third constraint is not satisfied with equality, which is a contradiction. Thus it has to be the case that $\theta_3(\boldsymbol{\mu}) = 1$. But this reduces the Segmentation 4 in (A43) to Segmentation 1 in (A19), which is already covered in region A.

We have shown that we can reduce the maximization problem subject to $A \cup B \cup C$ to a maximization problem subject to $A \cup \overline{B} \cup \overline{C}$, whose maximizer is given in equation (A42). Next we will find the optimal solution subject to region E.

Let us look at region $E \cap U_h$. Since for all $(\mu_h, \mu_l) \in E \cap U_h$, $\mu_h \geqslant \mu_l$, it has to be the case that $p_h(\boldsymbol{\mu}) \leqslant p_l(\boldsymbol{\mu})$. And also note that since $E \cap U_h \subset LF^+ \cap HF^-$, $p_h(\boldsymbol{\mu}) \geqslant c$. Giving this pricing scheme, we have the following segmentation, which we call "Segmentation 5"

$$\overset{l}{[-1, \theta_1(\boldsymbol{\mu})]}, \overset{h}{[\theta_1(\boldsymbol{\mu}), \theta_2(\boldsymbol{\mu})]}, \overset{F,l}{[\theta_2(\boldsymbol{\mu}), 0]}, \overset{F,h}{[0, \theta_3(\boldsymbol{\mu})]}, \overset{h}{[\theta_3(\boldsymbol{\mu}), 1]} \tag{A46}$$

where $-1 \leqslant \theta_1(\boldsymbol{\mu}) \leqslant \theta_2(\boldsymbol{\mu}) \leqslant 0 \leqslant \theta_3(\boldsymbol{\mu}) \leqslant 1$ are given by the following incentive constraints

$$\begin{aligned} 4\theta_1(\boldsymbol{\mu})\left(-\varepsilon - \frac{1}{2}(\mu_h - \mu_l)(\gamma - \varepsilon)\right) &\geqslant p_l(\boldsymbol{\mu}) - p_h(\boldsymbol{\mu}) \\ -4\theta_1(\boldsymbol{\mu})\frac{1}{2}\mu_h(\gamma - \varepsilon) &\geqslant p_l(\boldsymbol{\mu}) - c \\ 4\theta_2(\boldsymbol{\mu})\left(\varepsilon - \frac{1}{2}\mu_l(\gamma - \varepsilon)\right) &\leqslant p_h(\boldsymbol{\mu}) - c \\ 4\theta_3(\boldsymbol{\mu})\frac{1}{2}\mu_l(\gamma - \varepsilon) &= p_h(\boldsymbol{\mu}) - c \end{aligned} \tag{A47}$$

At least one of the first two inequalities hold with equality. The first constraint holds with equality if and only if the third constraint holds with equality. And if $\theta_1(\boldsymbol{\mu}) < \theta_2(\boldsymbol{\mu})$ the second constraint holds with inequality and the first and third constraints holds with equality. Now suppose $\theta_1(\boldsymbol{\mu}) < \theta_2(\boldsymbol{\mu})$. This means both the first and third constraints hold with equality. Given this structure $\mu_h = \frac{1}{2}(\theta_2(\boldsymbol{\mu}) - \theta_1(\boldsymbol{\mu}) + 1 - \theta_3(\boldsymbol{\mu}))$ and $\mu_l = \frac{\theta_1(\boldsymbol{\mu})+1}{2}$. We can also write $\theta_1(\boldsymbol{\mu}) = 2\mu_l - 1$ and $\theta_3(\boldsymbol{\mu}) = \theta_2(\boldsymbol{\mu}) + 2 - 2\mu_l - 2\mu_h$. From the last expression on $\theta_3(\boldsymbol{\mu})$ and the third and fourth constraints we can solve for $\theta_2(\boldsymbol{\mu})$ as $\theta_2(\boldsymbol{\mu}) = (-1 + \mu_l + \mu_h)\mu_l\frac{\gamma-\varepsilon}{\mu_l(\gamma-\varepsilon)-\varepsilon}$. A careful comparison of $\theta_1(\boldsymbol{\mu})$ and $\theta_2(\boldsymbol{\mu})$ reveals that $\theta_1(\boldsymbol{\mu}) - \theta_2(\boldsymbol{\mu}) < 0$ if and only if $\mu_h > \mu_l + \frac{\varepsilon(1-2\mu_l)}{\mu_l(\gamma-\varepsilon)}$. This means for $(\mu_h, \mu_l) \in \underline{E} = \left\{(\mu_h, \mu_l) \in E \cap U_h \mid \mu_h \leqslant \mu_l + \frac{\varepsilon(1-2\mu_l)}{\mu_l(\gamma-\varepsilon)}\right\}$, $\theta_1(\boldsymbol{\mu}) = \theta_2(\boldsymbol{\mu})$ and hence Segmentation 5 reduces to Segmentation 1, to which we know the optimizer is $\left(\mu_h^{*1}, \mu_l^{*1}\right) = \left(\frac{1}{3}, \frac{1}{3}\right) \in A$.

Now let us focus on $(\mu_h, \mu_l) \in \overline{E} = \left\{(\mu_h, \mu_l) \in E \cap U_h \mid \mu_h > \mu_l + \frac{\varepsilon(1-2\mu_l)}{\mu_l(\gamma-\varepsilon)}\right\}$ which we know generates Segmentation 5 with $\theta_1(\boldsymbol{\mu}) < \theta_2(\boldsymbol{\mu})$. It is the first, third and forth inequalities that hold with equality. Using the constraints and applying the substitution

$$z = \mu_l + \mu_h \tag{A48}$$

we can write the reduced form profits as

$$\widetilde{\pi}_M(z,\mu_l) = 2z(1-z)\mu_l(\gamma-\varepsilon)\frac{\mu_l(\gamma-\varepsilon)-2\varepsilon}{\mu_l(\gamma-\varepsilon)-\varepsilon} + 4\mu_l(1-2\mu_l)\left(\frac{1}{2}(\gamma-\varepsilon)(z-2\mu_l)+\varepsilon\right) \tag{A49}$$

where $\frac{2}{\rho-1} \leqslant \mu_l \leqslant \frac{1}{2}$ and $.2\mu_l \leqslant 2\mu_l + \frac{(1-2\mu_l)}{\mu_l(\rho-1)} \leqslant z \leqslant 1$.

Let us first look at how the profit changes with respect to z:

$$\begin{aligned} \frac{\partial\widetilde{\pi}_M(z,\mu_l)}{\partial z} &= 2(1-2z)\mu_l(\gamma-\varepsilon)\frac{\mu_l(\gamma-\varepsilon)-2\varepsilon}{\mu_l(\gamma-\varepsilon)-\varepsilon} + 2\mu_l(1-2\mu_l)(\gamma-\varepsilon) \qquad \text{(A50)} \\ &\geqslant \frac{1}{2}\frac{\varepsilon\mu_l(\rho-1)}{\mu_l(\rho-1)-1}\left(-2\mu_l^2(\rho-1)+2\mu_l+1\right) \end{aligned}$$

where, the inequality is due to $z \leqslant 1$ and $\frac{\partial^2\widetilde{\pi}_M(z,\mu_l)}{\partial z^2} \leqslant 0$.

The expression inside the parenthesis has roots $\mu_{l1} = \frac{1-\sqrt{2\rho-1}}{2(\rho-1)}$ and $\mu_{l2} = \frac{1+\sqrt{2\rho-1}}{2(\rho-1)}$. Observe that $\mu_{l1} < \frac{2}{\rho-1} < \mu_{l2}$ for $\rho > 5$. But this means $\frac{\partial\widetilde{\pi}_M(z,\mu_l)}{\partial z} > 0$ for all $(\mu_l, z) \in E \cap U_h$ such that $\mu_l < \mu_{l2}$ and $z < 1$ and hence

$$\arg \max_{2\mu_l+\frac{(1-2\mu_l)}{\mu_l(\rho-1)} \leqslant z \leqslant 1} \widetilde{\pi}_M(z,\mu_l) = 1 \text{ for all } \mu_l \leqslant \frac{1+\sqrt{2\rho-1}}{2(\rho-1)} \tag{A51}$$

Let us first focus on $\mu_l \leqslant \frac{1+\sqrt{2\rho-1}}{2(\rho-1)}$, plug $z = 1$ back into the profit function and try to find the maximizing μ_l.

$$\widetilde{\pi}_M(1,\mu_l) = \mu_l(1-2\mu_l)\left(\frac{1}{2}(\gamma-\varepsilon)(1-2\mu_l)+\varepsilon\right) \tag{A52}$$

The first derivative with respect to μ_l is

$$\frac{d\widetilde{\pi}_M(1,\mu_l)}{d\mu_l} = 6(\gamma-\varepsilon)\mu_l^2 - 4\gamma\mu_l + \frac{1}{2}(\gamma+\varepsilon) \tag{A53}$$

Looking at the roots of this first order derivative and analyzing the second order derivative reveal the maximizer as

$$\mu_l^{*5f} = \arg\max_{\frac{2}{(\rho-1)}\leqslant\mu_l\leqslant\frac{1+\sqrt{2\rho-1}}{2(\rho-1)}} \widetilde{\pi}_M(1,\mu_l) = \begin{cases} \frac{2}{\rho-1} \text{ if } \rho\leqslant 8+\sqrt{17} \\ \frac{2\rho-\sqrt{\rho^2+3}}{6(\rho-1)} \text{ if } 8+\sqrt{17}<\rho\leqslant\widehat{\rho}_2 \\ \frac{1+\sqrt{2\rho-1}}{2(\rho-1)} \text{ if } \rho>\widehat{\rho}_2 \end{cases} \tag{A54}$$

where $\widehat{\rho}_2 \sim 23.34$. From now on we can consider $\left(1-\mu_l^{*5f},\mu_l^{*5f}\right)$ as Segmentation 5 full market share maximizer.

Also notice that since $z \geqslant .2\mu_l$

$$\frac{\partial\widetilde{\pi}_M(z,\mu_l)}{\partial z} \leqslant \frac{2\varepsilon\mu_l(\rho-1)}{(\mu_l(\rho-1)-1)}\left(-6(\rho-1)\mu_l^2+2(\rho+4)\mu_l-3\right) \tag{A55}$$

The roots of this first order derivative and the second order derivative imply that

$$\arg\max_{2\mu_l\leqslant z\leqslant 1}\widetilde{\pi}_M(z,\mu_l) = 2\mu_l \text{ for all } \mu_l \geqslant \frac{\rho+4+\sqrt{\rho^2-10\rho+34}}{6(\rho-1)} \tag{A56}$$

When we plug this back into the profit function in (A49) to get $\widetilde{\pi}_M(2\mu_l,\mu_l)$, it is easily seen that $\frac{d\widetilde{\pi}_M(2\mu_l,\mu_l)}{d\mu_l} < 0$ for all $\mu_l \geqslant \frac{\rho+4+\sqrt{\rho^2-10\rho+34}}{6(\rho-1)}$ and for all $\rho > 5$. Hence

$$\arg\max_{\mu_l\geqslant\frac{\rho+4+\sqrt{\rho^2-10\rho+34}}{6(\rho-1)}}\widetilde{\pi}_M(2\mu_l,\mu_l) = \frac{\rho+4+\sqrt{\rho^2-10\rho+34}}{6(\rho-1)} \tag{A57}$$

Now let us consider $\frac{\rho+4+\sqrt{\rho^2-10\rho+34}}{6(\rho-1)} \geqslant \mu_l \geqslant \frac{1+\sqrt{2\rho-1}}{2(\rho-1)}$, which is the only region in which we might have an interior solution. Using the first order derivative we took in equation (A50) before, we can say that for all $\mu_l > \frac{1+\sqrt{2\rho-1}}{2(\rho-1)}$,

$$z(\mu_l) \equiv \arg \max_{1 \geq z \geqslant 2\mu_l + \frac{(1-2\mu_l)}{\mu_l(\rho-1)}} \widetilde{\pi}_M(z, \mu_l) = \frac{1}{2} + \frac{(1-2\mu_l)((\rho-1)\mu_l - 1)}{2((\rho-1)\mu_l - 2)} \tag{A58}$$

We plug this into the profit function and derive $\widetilde{\pi}_M(z(\mu_l), \mu_l)$. Then we take the first derivative with respect to μ_l and find that $\left.\frac{d\widetilde{\pi}_M(z(\mu_l),\mu_l)}{d\mu_l}\right|_{\mu_l = \frac{1+\sqrt{2\rho-1}}{2(\rho-1)}} \gtrless 0$ as $\rho \gtrless \widehat{\rho}_2 \sim 23.34$ and $\left.\frac{d\widetilde{\pi}_M(z(\mu_l),\mu_l)}{d\mu_l}\right|_{\mu_l = \frac{\rho+4+\sqrt{\rho^2-10\rho+34}}{6(\rho-1)}} < 0$ for all $\rho > 5$. These two respectively imply that for all $\rho > \widehat{\rho}_2$,

$$\arg \max_{\frac{\rho+4+\sqrt{\rho^2-10\rho+34}}{6(\rho-1)} \geqslant \mu_l \geqslant \frac{1+\sqrt{2\rho-1}}{2(\rho-1)}} \widetilde{\pi}_M(z(\mu_l), \mu_l) > \frac{1+\sqrt{2\rho-1}}{2(\rho-1)}$$

and for all $\rho > 5$,

$$\arg \max_{\frac{\rho+4+\sqrt{\rho^2-10\rho+34}}{6(\rho-1)} \geqslant \mu_l \geqslant \frac{1+\sqrt{2\rho-1}}{2(\rho-1)}} \widetilde{\pi}_M(z(\mu_l), \mu_l) < \frac{\rho+4+\sqrt{\rho^2-10\rho+34}}{6(\rho-1)}$$

And also careful algebraic analysis shows that when $\rho < \widehat{\rho}_2$, $\frac{d\widetilde{\pi}_M(z(\mu_l),\mu_l)}{d\mu_l} < 0$ for all $\frac{1+\sqrt{2\rho-1}}{2(\rho-1)} \leqslant \mu_l \leqslant \frac{\rho+4+\sqrt{\rho^2-10\rho+34}}{6(\rho-1)}$. We also showed above that $\frac{1+\sqrt{2\rho-1}}{2(\rho-1)} \lesseqgtr \frac{2\rho-\sqrt{\rho^2+3}}{6(\rho-1)} = \arg\max_{\mu_l} \widetilde{\pi}_M(1, \mu_l)$ as $\rho \gtrless \widehat{\rho}_2$. But these together with the Segmentation 5 full market share maximizer we found in (A54) reveal the Segmentation 5 maximizer as

$$\left(z^{*5}, \mu_l^{*5}\right) = \begin{cases} \left(1, \frac{2}{\rho-1}\right) \text{ if } \rho \leqslant 8+\sqrt{17} \\ \left(1, \frac{2\rho-\sqrt{\rho^2+3}}{6(\rho-1)}\right) \text{ if } 8+\sqrt{17} < \rho \leqslant \widehat{\rho}_2 \\ (z(\mu_l), \mu_l) \text{ for some } \mu_l > \frac{1+\sqrt{2\rho-1}}{2(\rho-1)} \text{if } \rho > \widehat{\rho}_2 \end{cases} \tag{A59}$$

Remember that we are still in CASE 3, i.e. $5 < \rho \leqslant 7$, hence the maximizer for region E

is given by $\left(z^{*5},\mu_l^{*5}\right)=\left(1,\frac{2}{\rho-1}\right)$. Now we should compare this maximum to the maximum of regions $A\cup B\cup C$ given in equation (A42) .Since $(\mu_h,\mu_l)=\left(1-\frac{2}{\rho-1},\frac{2}{\rho-1}\right)\in A\cup B\cup C$ it can be concluded that for $5<\rho\leqslant 7$, the maximizer over the whole region is

$$\arg\max_{(\mu_h,\mu_l)\in U}\pi_M(\mu_h,\mu_l)=\begin{cases}\left(\frac{1}{3},\frac{1}{3}\right) \text{ if } \rho\leqslant\widehat{\rho}_1\\ \left(\frac{1}{2}-\frac{2-\sqrt{3\rho^2-12\rho+13}}{6(\rho-1)},\frac{1}{2}+\frac{2-\sqrt{3\rho^2-12\rho+13}}{6(\rho-1)}\right) \text{ if } \widehat{\rho}_1\leqslant\rho\leqslant 7\end{cases} \tag{A60}$$

Letting $\rho_1=\widehat{\rho}_1$ proves proposition 4. ∎

Proof of Propositions 1.3-1.6, [Case 4.] [CASE 4: $\rho>7$] When $\rho>7$, all regions are non-empty. Let us go through regions A,B,C,E. We know that in region $A\cup\underline{B}\cup\underline{C}\cup\underline{E}$ the relevant solution is of Segmentation 1 with a maximum at $\left(\mu_h^{*1},\mu_l^{*1}\right)=\left(\frac{1}{3},\frac{1}{3}\right)\in\underline{E}$ for $\rho>7$.

In $\overline{B}\cup\overline{C}$, we know that the relevant solution is of Segmentation 2. Notice that when $\rho>7$, $\mu_h>\frac{2}{3}$ for all μ_l such that $(\mu_h,\mu_l)\in\overline{B}$. But using (A32) and (A38) and the fact that (A40) is increasing for $\rho>\frac{19}{3}$, we can deduce that for $\rho>7$, the Segmentation 2 maximizer is given by $\left(\mu_h^{*2f},1-\mu_h^{*2f}\right)$ where μ_h^{*2f} is defined in (A38). So far, we have found the maximums of segmentations 1 and 2 for $\rho>7$.

Before looking into region $\overline{E}$, which has a Segmentation 5 solution, let us try to analyze region D. Since for all $(\mu_h,\mu_l)\in D\cap U_h$, $\mu_h\geqslant\mu_l$, it has to be the case that $p_h(\boldsymbol{\mu})\leqslant p_l(\boldsymbol{\mu})$. And also note that since $D\subset LF^+\cap HF^-$, $p_h\geqslant c$. Giving this pricing scheme, we have the following segmentation, which we call "Segmentation 6"

$$[-1,\overset{l}{\theta_1}(\boldsymbol{\mu})],[\theta_1(\boldsymbol{\mu})\overset{h}{,}\theta_2(\boldsymbol{\mu})],[\theta_2\overset{F,l}{(\boldsymbol{\mu})},0],[0,\overset{F,h}{\theta_3}(\boldsymbol{\mu})],[\theta_3(\boldsymbol{\mu})\overset{h}{,}\theta_4(\boldsymbol{\mu})],[\theta_4\overset{l}{(\boldsymbol{\mu})},1] \tag{A61}$$

where $-1 \leqslant \theta_1(\boldsymbol{\mu}) \leqslant \theta_2(\boldsymbol{\mu}) \leqslant 0 \leqslant \theta_3(\boldsymbol{\mu}) \leqslant \theta_4(\boldsymbol{\mu}) \leqslant 1$ are given by the following incentive constraints

$$\begin{aligned}
4\theta_1(\boldsymbol{\mu})\left(-\varepsilon - \frac{1}{2}(\mu_h - \mu_l)(\gamma - \varepsilon)\right) &\geqslant p_l(\boldsymbol{\mu}) - p_h(\boldsymbol{\mu}) \\
-4\theta_1(\boldsymbol{\mu})\frac{1}{2}\mu_h(\gamma - \varepsilon) &\geqslant p_l(\boldsymbol{\mu}) - c \\
4\theta_2(\boldsymbol{\mu})\left(\varepsilon - \frac{1}{2}\mu_l(\gamma - \varepsilon)\right) &\leqslant p_h(\boldsymbol{\mu}) - c \\
4\theta_3(\boldsymbol{\mu})\frac{1}{2}\mu_l(\gamma - \varepsilon) &= p_h(\boldsymbol{\mu}) - c \\
4\theta_4(\boldsymbol{\mu})\left(\frac{1}{2}(\mu_h - \mu_l)(\gamma - \varepsilon) - \varepsilon\right) &\leqslant p_l(\boldsymbol{\mu}) - p_h(\boldsymbol{\mu})
\end{aligned} \tag{A62}$$

The argument about the first three constraints is the same as in what we discussed for region $\overline{E}$ in the previous case. Regarding the fifth constraint, if $\theta_4(\boldsymbol{\mu}) < 1$, it holds with equality. Given this structure $\mu_h = \frac{1}{2}(\theta_2(\boldsymbol{\mu}) - \theta_1(\boldsymbol{\mu}) + \theta_4(\boldsymbol{\mu}) - \theta_3(\boldsymbol{\mu}))$ and $\mu_l = \frac{\theta_1(\mu) - \theta_4(\mu)}{2} + 1$. Now suppose the second constraint holds with equality and also $\theta_4(\boldsymbol{\mu}) < 1$.This means $\theta_1(\boldsymbol{\mu}) = \theta_2(\boldsymbol{\mu})$. As we argued before, this implies the first and third constraints are slack and $\mu_h = \frac{\theta_4(\mu) - \theta_3(\mu)}{2}$ and $\mu_l = \frac{\theta_1(\mu) - \theta_4(\mu)}{2} + 1$. We can also write $\theta_1(\boldsymbol{\mu}) = 2\mu_l + \theta_4(\boldsymbol{\mu}) - 2$ and $\theta_3(\boldsymbol{\mu}) = \theta_4(\boldsymbol{\mu}) - 2\mu_h$. From the second and fourth constraints in (A62) we can derive

$$p_l(\boldsymbol{\mu}) - p_h(\boldsymbol{\mu}) = -2(\gamma - \varepsilon)(\theta_4(\boldsymbol{\mu})(\mu_h + \mu_l) - 2\mu_h) \tag{A63}$$

Using this price difference we can rewrite the LHS-RHS of the last constraint as

$$4(\theta_4(\boldsymbol{\mu})(\mu_h(\gamma - \varepsilon) - \varepsilon) - (\gamma - \varepsilon)\mu_h) < -4\varepsilon \leqslant 0$$

where the first inequality is due to $\mu_h \geqslant \frac{2}{\rho - 1}$ for all $(\mu_l, \mu_h) \in D$ and $\theta_4(\boldsymbol{\mu}) < 1$. Therefore the last constraint is slack, but this contradicts $\theta_4(\boldsymbol{\mu}) < 1$. Hence it can not be the case that both the second and the last constraint hold with equality. But this means that the necessary conditions for $\theta_1(\boldsymbol{\mu}) < \theta_2(\boldsymbol{\mu})$ are the same as we found in the previous case for Segmentation 5.

Suppose $\theta_4(\boldsymbol{\mu}) < 1$. We have just shown this implies all the constraints except the second one hold with equality and the second one is slack. We can write $\theta_3(\boldsymbol{\mu}) = \theta_2(\boldsymbol{\mu}) - 2\mu_h + \theta_4(\boldsymbol{\mu}) - \theta_1(\boldsymbol{\mu})$ and $\theta_1(\boldsymbol{\mu}) = 2\mu_l + \theta_4(\boldsymbol{\mu}) - 2$. Substituting in for $p_l(\boldsymbol{\mu}) - p_h(\boldsymbol{\mu})$ from the first constraint the LHS-RHS of the last constraint is

$$\begin{aligned}
& 4\theta_4(\boldsymbol{\mu})\left(\frac{1}{2}(\mu_h - \mu_l)(\gamma - \varepsilon) - \varepsilon\right) - 4\left(2\mu_l + \theta_4(\boldsymbol{\mu}) - 2\right)\left(-\varepsilon - \frac{1}{2}(\mu_h - \mu_l)(\gamma - \varepsilon)\right) \\
< \;& 4\left(\mu_h\mu_l(\gamma - \varepsilon) - \mu_l(\gamma - \varepsilon) + (1 - \mu_l)\left(\mu_l(\gamma - \varepsilon) - 2\varepsilon\right)\right)
\end{aligned}$$

where the inequality is due to $\theta_4(\boldsymbol{\mu}) < 1$ and rearranging. Notice that the last expression is strictly negative for all $\mu_h < \mu_l + \frac{2(1-\mu_l)}{\mu_l(\rho-1)}$ and hence the constraint is slack which contradicts $\theta_4(\boldsymbol{\mu}) < 1$. Therefore, if $\mu_h < \mu_l + \frac{2(1-\mu_l)}{\mu_l(\rho-1)}$, $\theta_4(\boldsymbol{\mu}) = 1$. But this reduces Segmentation 6 to Segmentation 5, which we analyzed in region $\overline{E}$ in the previous case. Let $D_1 = \left\{(\mu_l, \mu_h) \in D \mid \mu_h < \mu_l + \frac{2(1-\mu_l)}{\mu_l(\rho-1)}\right\}$ and $D_2 = \left\{(\mu_l, \mu_h) \in D \mid \mu_h \geqslant \mu_l + \frac{2(1-\mu_l)}{\mu_l(\rho-1)}\right\}$. We have shown that in region D_1 Segmentation 5 is valid whereas in region D_2 Segmentation 6 is valid.

Let us elaborate more on D_2. Notice that since $\mu_h \leqslant 1 - \mu_l$, any $(\mu_l, \mu_h) \in D_2$ has to satisfy $\mu_l + \frac{2(1-\mu_l)}{\mu_l(\rho-1)} \leqslant 1 - \mu_l$. But this is never satisfied and hence $D_2 = \emptyset$ for $7 < \rho < 7 + 4\sqrt{2}$. And for $\rho \geqslant 7 + 4\sqrt{2}$, it implies that $\frac{\rho+1-\sqrt{\rho^2-14\rho+17}}{4(\rho-1)} \leqslant \mu_l \leqslant \frac{\rho+1+\sqrt{\rho^2-14\rho+17}}{4(\rho-1)}$.

Now we can combine $\overline{E}$ and D_1 and apply the Segmentation 5 analysis we did in the previous case to these regions. As we found in (A59), for $\rho \leqslant 8+\sqrt{17}$ the maximizer subject to regions $\overline{E}\cup D_1$ is $(z,\mu_l)=\left(1,\frac{2}{\rho-1}\right)$. When we analyze the case where $\rho > 8+\sqrt{17}$ referring to (A59), simple algebra shows that $\frac{2\rho-\sqrt{\rho^2+3}}{6(\rho-1)} \lesseqgtr \frac{\rho+1-\sqrt{\rho^2-14\rho+17}}{4(\rho-1)}$ as $\rho \lesseqgtr \frac{25}{3}+\frac{8}{3}\sqrt{7}$. Thus we can say that the maximizer subject to regions $\overline{E}\cup D_1$ is $(z,\mu_l)=\left(1,\frac{2\rho-\sqrt{\rho^2+3}}{6(\rho-1)}\right)$ if $8+\sqrt{17}\leqslant\rho\leqslant\frac{25}{3}+\frac{8}{3}\sqrt{7}$. Now before going on with the maximization in regions $\overline{E}\cup D_1$, since region D_2 interferes with regions $\overline{E}\cup D_1$, we will include D_2 into the analysis and try to find the maximum of $\overline{E}\cup D_1\cup D_2$.

We start by writing the profit function for region D_2, using the incentive constraints we introduced for Segmentation 6

$$\begin{aligned}\widetilde{\pi}_M(z,\mu_l) &= 2z(1-z)\mu_l(\gamma-\varepsilon)\frac{\mu_l(\gamma-\varepsilon)-2\varepsilon}{\mu_l(\gamma-\varepsilon)-\varepsilon} \\ &\quad +4\mu_l(1-\mu_l)\left(\frac{1}{2}(\gamma-\varepsilon)(z-2\mu_l)+\varepsilon\right)\left(1-\frac{2\varepsilon}{(\gamma-\varepsilon)(z-2\mu_l)}\right)\end{aligned} \tag{A64}$$

Let us first find the full market share maximizer in region D_2. We can write the full market share profit as

$$\widetilde{\pi}_M(1,\mu_l)=\mu_l(1-\mu_l)\left(\frac{1}{2}(\gamma-\varepsilon)(1-2\mu_l)+\varepsilon\right)\left(1-\frac{2\varepsilon}{(\gamma-\varepsilon)(1-2\mu_l)}\right) \tag{A65}$$

And taking the derivative with respect to μ_l gives the following solution

$$\arg\max_{\frac{\rho+1-\sqrt{\rho^2-14\rho+17}}{4(\rho-1)}\leqslant\mu_l\leqslant\frac{\rho+1+\sqrt{\rho^2-14\rho+17}}{4(\rho-1)}}\widetilde{\pi}_M(1,\mu_l)=\begin{cases}\frac{\rho+1-\sqrt{\rho^2-14\rho+17}}{4(\rho-1)} \\ \text{if } \rho\leqslant 5+2\sqrt{2}+2\sqrt{5+4\sqrt{2}} \\ \mu(\rho) \text{ otherwise}\end{cases} \tag{A66}$$

where $\mu(\rho) = \frac{1}{2} - \frac{\sqrt{(\rho-1)^2(5+(\rho-2)\rho+\sqrt{73+\rho(\rho-2)(58+(\rho-2)\rho)})}}{2\sqrt{6}(\rho-1)^2}$.

Comparing this to the full market share maximum of regions $\overline{E}\cup D_1$,we get the full market share maximizer in $\overline{E}\cup D$ to be $\left(1-\mu_l^{*56f}, \mu_l^{*56f}\right)$, where

$$\mu_l^{*56f} = \begin{cases} \frac{2}{\rho-1} \text{ if } \rho \leqslant 8+\sqrt{17} \\ \frac{2\rho-\sqrt{\rho^2+3}}{6(\rho-1)} \text{ if } 8+\sqrt{17} < \rho \leqslant \widehat{\rho}_3 \\ \mu_l(\rho) \text{ otherwise} \end{cases} \tag{A67}$$

We can interpret $\left(1-\mu_l^{*56f}, \mu_l^{*56f}\right)$ as the segmentations 5 and 6 full market share maximizer.

Before finding the interior maximum, simple algebra shows $\pi_M\left(1-\mu_l^{*56f}, \mu_l^{*56f}\right) > \pi_M\left(\frac{1}{3}, \frac{1}{3}\right)$. Therefore, we can eliminate Segmentation 1 for $\rho > 7$. Recall that we have also found the maximum of Segmentation 2 is $\left(\mu_h^{*2f}, 1-\mu_h^{*2f}\right)$ given by (A38). Comparing it with $\left(1-\mu_l^{*56f}, \mu_l^{*56f}\right)$ we see that for $\rho \leqslant \frac{11}{2}+\frac{1}{2}\sqrt{17}$ $\left(\mu_h^{*2f}, 1-\mu_h^{*2f}\right)$ performs better. Thus, now, all that remains is finding the interior maximum of segmentations 5 and 6 and comparing them to their full market share maximums.

We already showed in the previous case that the maximizer of Segmentation 5 becomes interior for $\rho > \widehat{\rho}_2$. And the interior solution has the form $(z(\mu_l), \mu_l)$. However, careful analysis shows that $(z(\mu_l), \mu_l) \in D_2$. Using (A64) and (A65), it can easily be shown that $(z(\mu_l), \mu_l)$ gives a higher profit if it can be supported by Segmentation 6 than when it belongs to Segmentation 5. But this means for $\rho > \widehat{\rho}_2$,

$$\max_{(\mu_h,\mu_l)\in\overline{E}\cup D_1} \pi_M(\mu_h, \mu_l) < \max_{(\mu_h,\mu_l)\in D_2} \pi_M(\mu_h, \mu_l)$$

Therefore for $\rho > \widehat{\rho}_2$, we can forget about Segmentation 5 and focus on the interior maximum for Segmentation 6. We know the full market share maximizer in region D_2. Careful computations show that, $\left[\frac{\partial \widetilde{\pi}_M(z,\mu_l)}{\partial z}\right]_{z=1,\mu_l=\mu_l(\rho)} \lesseqgtr 0$ as $\rho \gtreqless \widehat{\rho}_4 \sim 30.05$. This means that for $\rho > \widehat{\rho}_4$, the maximum will be attained at $z < 1$, since $\widetilde{\pi}_M(z,\mu_l)$ increases as we decrease z even at the point where the full market share reaches its maximum. Notice also that at $\rho = \widehat{\rho}_4$, $(z,\mu_l) = (1,\mu_l(\rho))$ is the local maximum. Careful computations also show that for $\rho \leqslant \widehat{\rho}_4$ it continues to be the global maximum.

Combining all these we can write the maximizer for $\rho \geqslant 7$ as

$$(\mu_h^*,\mu_l^*) = \begin{cases} \left(\frac{1}{2} - \frac{2-\sqrt{3\rho^2-12\rho+13}}{6(\rho-1)}, \frac{1}{2} + \frac{2-\sqrt{3\rho^2-12\rho+13}}{6(\rho-1)}\right) \text{ if } \rho \leqslant \frac{11}{2} + \frac{1}{2}\sqrt{17} \\ \left(1, \frac{2}{\rho-1}\right) \text{ if } \rho \leqslant 8+\sqrt{17} \\ \left(1, \frac{2\rho-\sqrt{\rho^2+3}}{6(\rho-1)}\right) \text{ if } 8+\sqrt{17} < \rho \leqslant \widehat{\rho}_3 \\ (1-\mu(\rho), \mu(\rho)) \text{ if } \widehat{\rho}_3 < \rho \leqslant \widehat{\rho}_4 \\ \text{some } (\mu_h,\mu_l) \in D_2 \text{ such that } \mu_h+\mu_l < 1 \text{ if } \rho > \widehat{\rho}_4 \end{cases} \tag{A68}$$

Now letting $\rho_2 = 8+\sqrt{17}$, $\rho_3 = \widehat{\rho}_3$ and $\rho_4 = \widehat{\rho}_4$ proves Propositions 3-6.

∎

Proof of Proposition 1.7. For $\rho \leqslant \rho_1$, Proposition 1.4 gives the characterization for the unique equilibrium as $\mu_h(\mathbf{p}^*) = \mu_l(\mathbf{p}^*) = \frac{1}{3}$, hence $\mu_h(\mathbf{p}^*)$ and $\mu_l(\mathbf{p}^*)$ remain unchanged with γ and ε.

For all $\rho_1 < \rho \leqslant \rho_3$, the differentiation of each μ_h^* term in equation (A68) directly reveal that $\frac{d\mu_h^*}{d\rho} > 0$ and therefore

$$d\left(\frac{\mu_h^*}{\mu_l^*}\right) = d\left(\frac{\mu_h^*}{1-\mu_h^*}\right) = \frac{1}{\left(1-\mu_h^*\right)^2}\frac{d\mu_h^*}{d\rho} > 0 \tag{A69}$$

which means the degree of unbalance is increasing in ρ.

For $\rho_3 \leqslant \rho < \rho_4$, the equilibrium as we proved in the proof of proposition 5 is given by $\mu_h(\mathbf{p}^*) = 1 - \mu_l(\rho)$ and $\mu_l(\mathbf{p}^*) = \mu(\rho)$, where $\mu(\rho)$ is defined in (A66). Here simple algebra shows that for all $\rho > 7$, $\frac{d\mu_l(\rho)}{d\rho} > 0$.and hence $d\left(\frac{\mu_h^*}{\mu_l^*}\right) < 0$.

To find the change in degree of shuffling, first notice that the degree of shuffling is given by

$$\frac{1-\theta_4(\boldsymbol{\mu})}{\theta_1(\boldsymbol{\mu})+1} \tag{A70}$$

where $\theta_4(\boldsymbol{\mu})$ and $\theta_1(\boldsymbol{\mu})$ are as described in (A62). Notice that Form 6 distribution implies $\theta_1(\boldsymbol{\mu}) = 2\mu_l + \theta_4(\boldsymbol{\mu}) - 2$. Using this and the incentive constraints in (A62) we can derive

$$\theta_4(\boldsymbol{\mu}) = \frac{(1-\mu_l)\left(\frac{1}{2}(\mu_h-\mu_l)(\gamma-\varepsilon)+\varepsilon\right)}{\frac{1}{2}(\mu_h-\mu_l)(\gamma-\varepsilon)} \tag{A71}$$

After substituting these and $\mu_l(\mathbf{p}^*) = \mu_l(\rho)$ back into (A70), simple algebra shows that $\frac{d}{d\rho}\left(\frac{1-\theta_4(\mu)}{\theta_1(\mu)+1}\right) > 0$. Thus the degree of shuffling is increasing.

Now let us prove point (4). We will first show that for $\rho > \rho_4$, at the local maximum, the total measure z is decreasing and μ_l is increasing. Let (z^*, μ_l^*) denote the optimal choices for $\rho > \rho_4$. By local optimality of (z^*, μ_l^*) we know that

$$\left.\frac{\partial\left(\frac{\tilde{\pi}_M(z,\mu_l)}{\varepsilon}\right)}{\partial z}\right|_{(z^*,\mu_l^*)} = \left.\frac{\partial\left(\frac{\tilde{\pi}_M(z,\mu_l)}{\varepsilon}\right)}{\partial \mu_l}\right|_{(z^*,\mu_l^*)} = 0$$

Simple algebra shows that this implies at (z^*, μ_l^*) $\frac{\partial^2\left(\frac{\tilde{\pi}_M(z,\mu_l)}{\varepsilon}\right)}{\partial z \partial z} \leqslant 0$, $\frac{\partial^2\left(\frac{\tilde{\pi}_M(z,\mu_l)}{\varepsilon}\right)}{\partial \mu_l \partial \mu_l} \leqslant 0$, $\frac{\partial^2\left(\frac{\tilde{\pi}_M(z,\mu_l)}{\varepsilon}\right)}{\partial z \partial \rho} \leqslant 0$ and $\frac{\partial^2\left(\frac{\tilde{\pi}_M(z,\mu_l)}{\varepsilon}\right)}{\partial \mu_l \partial \rho} \geqslant 0$. Notice that the local optimality of (z^*, μ_l^*) also implies that the Hessian matrix is negative semi-definite at (z^*, μ_l^*). Through the implicit function theorem, these second order conditions prove that $\frac{dz^*}{d\rho} < 0$ and $\frac{d\mu_l^*}{d\rho} > 0$.

We can write the unbalance level as $\frac{z^*-\mu_l^*}{\mu_l^*}$. Since z^* is decreasing and μ_l^* is increasing, it is clear that $\frac{d\left(\frac{z^*-\mu_l^*}{\mu_l^*}\right)}{d\rho} \leqslant 0$.

The only thing that remains to be shown is what happens to the degree of shuffling as $\rho \to \infty$. First let us find the limiting equilibria. From (A64) we can write the limit profits as $\varepsilon \to 0$ as

$$\widetilde{\pi}_M(z, \mu_l) = 2\gamma\mu_l\left(z(1-z) + (1-\mu_l)(z - 2\mu_l)\right) \tag{A72}$$

Given μ_l, the first and second order conditions yields the optimal z to be $1 - \frac{\mu_l}{2}$. Then we can write

$$\widetilde{\pi}_M\left(1 - \frac{\mu_l}{2}, \mu_l\right) = 2\gamma\mu_l\left(\left(1 - \frac{\mu_l}{2}\right)\frac{\mu_l}{2} + (1-\mu_l)\left(1 - \frac{5\mu_l}{2}\right)\right) \tag{A73}$$

$$\frac{\partial\widetilde{\pi}_M(z(\mu_l), \mu_l)}{\partial\mu_l} = \gamma\left(\frac{27}{2}\mu_l^2 - 12\mu_l + 2\right) \tag{A74}$$

The first order derivative and the second order conditions imply that the maximizing pair is $(\mu_h^*, \mu_l^*) = \left(\frac{2}{3}, \frac{2}{9}\right)$. The marginal buyer points can be derived easily from (A62). The degree of shuffling using (A62) is

$$\frac{1-\theta_4(\boldsymbol{\mu})}{\theta_1(\boldsymbol{\mu})+1} = \frac{(\mu_h - \mu_l)(\rho - 1)\mu_l - 2(1-\mu_l)}{(\mu_h - \mu_l)(\rho - 1)\mu_l + 2(1-\mu_l)} \tag{A75}$$

Since $\mu_h^* - \mu_l^* = \frac{4}{9} > 0$, it is easily seen that $\lim_{\rho\to\infty} \frac{1-\theta_4(\boldsymbol{\mu}^*)}{\theta_1(\boldsymbol{\mu}^*)+1} = 1$ ∎

Proof of Proposition 1.8. Point 1 is straight forward. To prove point 2, let us focus on the efficient allocation with $\mu_h^s > \mu_l^s$. From the proof of Proposition 1.1 we know that the efficient allocation being shuffled implies $\frac{3\rho-5-\sqrt{\rho^2-14\rho+17}}{4(\rho-1)} < \mu_h^s < \frac{3\rho-5+\sqrt{\rho^2-14\rho+17}}{4(\rho-1)}$. Also notice that due to full market share $\mu_l^s = 1 - \mu_h^s$. But these together imply that $\mu_l^s > \frac{\rho+1-\sqrt{\rho^2-14\rho+17}}{4(\rho-1)} > \frac{2}{(\rho-1)}$, where the second inequality can be derived by simple algebra. But this means that $(\mu_h^s, \mu_l^s) \in HF^-$, which implies that all $p_h \leqslant c$ captures full market. Hence the profit maximizing choice is $p_h^s = c$.

Using equation (A11) it could easily be derived that

$$\begin{aligned} & U_M\left(\theta_1^s\left(\mu_h^s\right), l, \mu_h^s\right) - U_M\left(\theta_1^s\left(\mu_h^s\right), h, 1-\mu_h^s\right) \\ = \; & U_M\left(\theta_1^s\left(\mu_h^s\right) + 2\mu_h^s, l, \mu_h^s\right) - U_M\left(\theta_1^s\left(\mu_h^s\right) + 2\mu_h^s, h, 1-\mu_h^s\right) \end{aligned} \tag{A76}$$

where $\theta_1^s\left(\mu_h^s\right)$ is given in (A11). But this means that given (μ_h^s, μ_l^s),

$$p_l^s - c = U_M\left(\theta_1^s\left(\mu_h^s\right), l, \mu_h^s\right) - U_M\left(\theta_1^s\left(\mu_h^s\right), h, 1-\mu_h^s\right) \tag{A77}$$

is the only price that makes the buyers with type $\theta_1^s\left(\mu_h^s\right)$ and $\theta_1^s\left(\mu_h^s\right) + 2\mu_h^s$ indifferent. By supermodularity, their indifference is a necessary and sufficient condition for the implementation of (μ_h^s, μ_l^s).

Also notice that (μ_h^s, μ_l^s) is the unique distribution that (p_h^s, p_l^s) implements given the refinement we introduced earlier. ∎

Proof of Proposition 1.9. We know from Proposition 1.1 that for all $\rho < \rho_s$ the efficient allocation is balanced. Proofs of propositions 1.1 and 1.3-1.6 revealed that $\rho_s > 23 > 7 > \rho_1$. This proves the first part of the proposition.

For all $\rho > \rho_1$, the equilibria imply unbalanced distributions. Together with points (1) and (2) in Proposition 1.7, this proves the second part for all $\rho < \rho_s$. Now let us try to prove the result for $\rho > \rho_s$. Notice the proofs of propositions 1.1 and 1.3-1.6 revealed that $\rho_s > 23 > 14.82 \sim \rho_3$. We also know that for all $\rho > \rho_3$, the degree of balance of the equilibrium distribution decreases continuously. We also know from Proposition 1.2 that for all $\rho > \rho_s$, the degree of balance of the efficient allocations increases continuously. This means the difference between the degrees of unbalance of equilibrium and efficient distributions decreases for $\rho > \rho_3$. Let us compare the degree of balance for both as $\rho \to \infty$. As $\rho \to \infty$, using point (5) in Proposition 1.7, the degree of balance of the equilibrium distribution is 3,whereas, using Proposition 1.2, the degree of balance of the efficient allocation is $\frac{2+\sqrt{7}}{3}$. Since $\frac{2+\sqrt{7}}{3} < 3$, this means the difference is always positive for $\rho > \rho_3$. The last point in the proposition follows directly from the fact that $\rho_s > \rho_3$ and the comparative statics of degree of unbalance given in Propositions 1.2 and 1.7. ∎

Proof of Proposition 1.10. In the proof of Proposition 1.9 we argued that $\rho_s > \rho_3$. It could also be shown numerically that $\rho_s < 30 < \rho_4$. Then points (1) and (2) can are direct results of Proposition 1.8 which reveals that $p_i^s(\gamma, \varepsilon) = c$ for all $i \in \{l, h\}$ and Propositions 3 and 4 which state equilibrium prices in comparison to c. Point (3) for can be derived the same way from Propositions 4 and 5 for $\rho < \rho_s$. However, if $\rho_s \leqslant \rho < \rho_4$, then a more careful analysis is necessary. Point (2) in Proposition 1.9 states that for this range of parameters the degree of unbalance is higher for equilibrium distributions than efficient distributions. However we also know that again for this range of parameters equilibrium involves full market share for seller M. Consider the equilibrium and efficient distribution where $\mu_h > \mu_l$ and let

μ_h^* and μ_h^s denote the equilibrium and efficient distributions respectively. First notice that $p_h^s(\gamma,\varepsilon) = p_h^*(\gamma,\varepsilon) = c$. The former facts imply that $\mu_h^* > \mu_h^s$. Let us compare the marginal buyer with a negative type for the seller, θ_1^* to the switch point buyer for efficiency, θ_1^E, which can respectively be derived from the incentive constraints in (A62) imposing $z = 1$ and (A11).

$$\theta_1^* - \theta_1^s = \left(-\mu_h^* + \frac{2\mu_h^*}{(2\mu_h^* - 1)(\rho - 1)}\right) - \left(-\mu_h^s + \frac{2\mu_h^s}{(2\mu_h^s - 1)(\rho - 1)}\right) < 0 \qquad \text{(A78)}$$

since $\mu_h^* > \mu_h^s$. Note that by Proposition 1.8, $p_l^s(\gamma,\varepsilon) = U_M(\theta_1^s, l, \mu_h^s) - U_M(\theta_1^s, h, \mu_l^s) + c = -4\theta_1^s\left(\varepsilon + \frac{1}{2}(2\mu_h^s - 1)(\gamma - \varepsilon)\right) + c$ and by the first incentive constraint in (A62) $p_l^*(\gamma,\varepsilon) = U_M(\theta_1^*, l, \mu_h^*) - U_M(\theta_1^*, h, \mu_l^*) + c = -4\theta_1^*\left(\varepsilon + \frac{1}{2}(2\mu_h^* - 1)(\gamma - \varepsilon)\right) + c$. Since $\theta_1^* < \theta_1^s < 0$ and $\mu_h^* > \mu_h^s$ it is evident that $p_l^*(\gamma,\varepsilon) > p_l^s(\gamma,\varepsilon)$.

The last part requires a more complicated analysis. From the proof of proposition 7 we know that $\frac{dz^*}{d\rho} < 0$ and $\frac{d\mu_l^*}{d\rho} > 0$. But this implies that $\frac{d(\mu_h^* - \mu_l^*)}{d\rho} < 0$. We also know from Proposition 1.2 and its proof that $\frac{d(\mu_h^s - \mu_l^s)}{d\rho} > 0$. By point (2) in Proposition 1.8 and by the fact that at $\rho = \rho_4$ both the efficient and the equilibrium distributions involve full market share, we can say that at $\rho = \rho_4$, $\mu_h^* - \mu_l^* > \mu_h^s - \mu_l^s$. We also know that at $\rho = \rho_4$ the optimum becomes an interior optimum and the equilibrium evolves continuously. From the last parts of Propositions 1.2 and 1.7 we can see that $\lim_{\rho\to\infty}(\mu_h^* - \mu_l^*) - (\mu_h^s - \mu_l^s) = \frac{4}{9} - \left(\frac{\sqrt{7}}{3} - \frac{2}{3}\right) = \frac{10 - 3\sqrt{7}}{9} > 0$. Hence we can say that $\mu_h^* - \mu_l^* > \mu_h^s - \mu_l^s$ for all $\rho \geqslant \rho_4$. The same way in can be shown that $\mu_h^* > \mu_h^s$ for all $\rho \geqslant \rho_4$. Now notice that $\theta_1^* = (1 - \mu_l^*)\left(-1 + \frac{2}{(\mu_h^* - \mu_l^*)(\rho - 1)}\right) < \mu_h^s\left(-1 + \frac{2}{(2\mu_h^s - 1)(\rho - 1)}\right) = \theta_1^s$ for all $\rho \geqslant \rho_4$, where the inequality is due to $\mu_h^* < 1 - \mu_l^*$, $\mu_h^* - \mu_l^* > 2\mu_h^* - 1$, the fact that the term in parenthesis in negative since $(\mu_h^*, \mu_l^*) \in D$ and

$\mu_h^* > \mu_h^s$. Notice that $p_l^*(\gamma,\varepsilon) - p_h^*(\gamma,\varepsilon) = -4\theta_1^*\left(\varepsilon + \frac{1}{2}(\mu_h^* - \mu_l^*)(\gamma-\varepsilon)\right)$ and $p_l^s(\gamma,\varepsilon) - p_h^s(\gamma,\varepsilon) = -4\theta_1^s\left(\varepsilon + \frac{1}{2}(\mu_h^s - \mu_h^s)(\gamma-\varepsilon)\right)$. Since $\theta_1^* < \theta_1^s < 0$ and $\mu_h^* - \mu_l^* > \mu_h^s - \mu_l^s$, $p_l^*(\gamma,\varepsilon) - p_h^*(\gamma,\varepsilon) > p_l^s(\gamma,\varepsilon) - p_h^s(\gamma,\varepsilon)$ for all $\rho \geqslant \rho_4$. ∎

1.8.2 Multiple Equilibria and Pareto Ranking

Let $(i^*(\theta,\mathbf{p},\widehat{\mu}_l,\widehat{\mu}_h), s^*(\theta,\mathbf{p},\widehat{\mu}_l,\widehat{\mu}_h))$ be type θ buyer's optimal choices given the prices and her belief, i.e.

$$\left(i^*(\theta,\mathbf{p},\widehat{\mu}_l,\widehat{\mu}_h), s^*(\theta,\mathbf{p},\widehat{\mu}_l,\widehat{\mu}_h)\right) = \arg\max_{i\in\{l,h\},s\in\{M,F\}} U\left(\theta,i,s,\widehat{\mu}_j,p_i\right)$$

Then let $(\widetilde{\mu}_l(\mathbf{p},\widehat{\mu}_l,\widehat{\mu}_h), \widetilde{\mu}_h(\mathbf{p},\widehat{\mu}_l,\widehat{\mu}_h))$ be the vector of measures implied by these optimal choices. Equilibrium requires consistency of beliefs, hence $(\mu_h(\mathbf{p}),\mu_l(\mathbf{p}))$ is an equilibrium if

$$\left(\widetilde{\mu}_l(\mathbf{p},\mu_h(\mathbf{p}),\mu_l(\mathbf{p})), \widetilde{\mu}_h(\mathbf{p},\mu_h(\mathbf{p}),\mu_l(\mathbf{p}))\right) = (\mu_h(\mathbf{p}),\mu_l(\mathbf{p}))$$

There may be multiple fixed points due to the coordination element inherent in the problem. Each $(\mu_h(\mathbf{p}),\mu_l(\mathbf{p}))$ represents an equilibrium market share of seller M for each product when he announces the prices as $\mathbf{p}$.

Lemma 1.7 (PARETO DOMINANCE).

The perfect Bayesian equilibria in the subgame following the announcement of any $\mathbf{p} \in \mathbf{R}^2$ *are strictly pareto ranked and an increase in the pareto rank coincides with a strict increase in the market share of seller* M *for both products.*

Proof of Lemma 1.7. Suppose $\varepsilon < \frac{1}{2}$ and $\gamma > 0$. Suppose there exists a price vector $\mathbf{p}$ that generates two-fixed points, i.e. there exists two non-zero vectors (μ_l, μ_h) and (μ_l', μ_h') such that $\mu_l' < \mu_l$ and $\mu_h' > \mu_h$ and

$$\begin{aligned} \left(\widetilde{\mu}_l\left(\mathbf{p},\mu_l,\mu_h\right),\widetilde{\mu}_h\left(\mathbf{p},\mu_l,\mu_h\right)\right) &= \left(\mu_l,\mu_h\right) \\ \left(\widetilde{\mu}_l\left(\mathbf{p},\mu_l',\mu_h'\right),\widetilde{\mu}_h\left(\mathbf{p},\mu_l',\mu_h'\right)\right) &= \left(\mu_l',\mu_h'\right) \end{aligned} \tag{B1}$$

But due to Lemma 1.1 and equation (4), this means that

$$\begin{aligned} U\left(\theta,h,M,\mu_l,p_h\right) &> U\left(\theta,h,M,\mu_l',p_h\right) \\ U\left(\theta,l,M,\mu_h,p_l\right) &< U\left(\theta,l,M,\mu_h',p_l\right) \end{aligned} \tag{B2}$$

Let $\Theta_L = \{\theta \in [-1,1] \mid (i^*(\theta,\mathbf{p},\mu_l,\mu_h), s^*(\theta,\mathbf{p},\mu_l,\mu_h)) = (l,M)\}$. The fact that these buyers choose to buy l from seller M when they believe in (μ_l,μ_h) implies that for all $\theta \in \Theta_L$,

$$U\left(\theta,h,M,\mu_l,p_h\right),U\left(\theta,l,F,\mu_h,p_l\right),U\left(\theta,h,F,\mu_h,p_l\right) \leqslant U\left(\theta,l,M,\mu_h,p_l\right) \tag{B3}$$

These combined with equation (B2) imply that

$$U\left(\theta,h,M,\mu_l',p_h\right),U\left(\theta,l,F,\mu_h,p_l\right),U\left(\theta,h,F,\mu_h,p_l\right) < U\left(\theta,l,M,\mu_h',p_l\right) \tag{B4}$$

But from (B4) we can deduce $(i^*(\theta,\mathbf{p},\mu_l',\mu_h'), s^*(\theta,\mathbf{p},\mu_l',\mu_h')) = (l,M)$ for all $\theta \in \Theta_L$. Thus, the buyers who used to buy product l with beliefs (μ_l,μ_h) are still buying product l with (μ_l',μ_h'). But this means $\widetilde{\mu}_l(\mathbf{p},\mu_l',\mu_h') > \widetilde{\mu}_l(\mathbf{p},\mu_l,\mu_h)$, which contradicts the fact that both (μ_l,μ_h) and (μ_l',μ_h') are fixed points as given by equation (B1). Thus if $\mu_l' < \mu_l$, it has to be the case that $\mu_h' \leqslant \mu_h$. So far we have shown that the fixed-points form a chain, i.e.

they can be ordered. (Note that if $\varepsilon = \frac{1}{2}$ or $\gamma = 0$, the problem is trivial and not of finding the fixed-point because $(\widetilde{\mu}_l(\mathbf{p}, \widehat{\mu}_l, \widehat{\mu}_h), \widetilde{\mu}_h(\mathbf{p}, \widehat{\mu}_l, \widehat{\mu}_h))$ is independent of $(\widehat{\mu}_l, \widehat{\mu}_h)$).

Now suppose there exists a price vector $\mathbf{p}$ that generates two-fixed points which are ordered as we found, i.e. there exists two non-zero vectors (μ_l, μ_h) and (μ_l', μ_h') such that $\mu_l' < \mu_l$ and $\mu_h' \leqslant \mu_h$. Notice that due to Lemma 1.1 and equation (4), since $\mu_l' < \mu_l$ and $\mu_h' \leqslant \mu_h$, for all $\theta \in [-1, 1]$

$$\begin{aligned} & U\left(\theta, i^*\left(\theta, \mathbf{p}, \mu_l', \mu_h'\right), s^*\left(\theta, \mathbf{p}, \mu_l', \mu_h'\right), \mu_{j^*}', p_{i^*}\right) \\ \leqslant\ & U\left(\theta, i^*\left(\theta, \mathbf{p}, \mu_l', \mu_h'\right), s^*\left(\theta, \mathbf{p}, \mu_l', \mu_h'\right), \mu_{j^*}, p_{i^*}\right) \end{aligned} \tag{B5}$$

But optimality of the buyers' choices $i^*(\theta, \mathbf{p}, \mu_l, \mu_h), s^*(\theta, \mathbf{p}, \mu_l, \mu_h)$ when their belief is (μ_l, μ_h) implies that for all $\theta \in [-1, 1]$

$$\begin{aligned} & U\left(\theta, i^*\left(\theta, \mathbf{p}, \mu_l', \mu_h'\right), s^*\left(\theta, \mathbf{p}, \mu_l', \mu_h'\right), \mu_{j^*}, p_{i^*}\right) \\ \leqslant\ & U\left(\theta, i^*\left(\theta, \mathbf{p}, \mu_l, \mu_h\right), s^*\left(\theta, \mathbf{p}, \mu_l, \mu_h\right), \mu_{j^*}, p_{i^*}\right) \end{aligned} \tag{B6}$$

Equations (B5) and (B6) can be combined to yield

$$\begin{aligned} & U\left(\theta, i^*\left(\theta, \mathbf{p}, \mu_l', \mu_h'\right), s^*\left(\theta, \mathbf{p}, \mu_l', \mu_h'\right), \mu_{j^*}', p_{i^*}\right) \\ \leqslant\ & U\left(\theta, i^*\left(\theta, \mathbf{p}, \mu_l, \mu_h\right), s^*\left(\theta, \mathbf{p}, \mu_l, \mu_h\right), \mu_{j^*}, p_{i^*}\right) \end{aligned} \tag{B7}$$

for all $\theta \in [-1, 1]$. Also notice that equation (B5) is a strict inequality for all $\theta \in [-1, 1]$ such that $(i^*(\theta, \mathbf{p}, \mu_l', \mu_h'), s^*(\theta, \mathbf{p}, \mu_l', \mu_h')) = (l, M)$. Thus for those buyers equation (B6) also holds with strict inequality, which proves that the fixed-points are pareto ranked. ∎

Chapter 2

Information Transmission to Multiple Audiences

2.1 Introduction

There are numerous occasions in real life where one informed party has to convey information to uninformed parties who afterwards undertake some actions that affect both the informed and uninformed parties. We would like to talk about occasions where there is more than one uninformed party (receivers of information) that the informed party (sender of information) has to pass information to. When we say "more than one uninformed party", we are referring to groups of uninformed parties where each party is identified with a different interest in the information to be received. The existence of more than one receiver creates different possibilities for the information transmission. In particular the sender can have private meetings with each receiver or he could organize a public meeting through which he addresses all the receivers. For example a politician could meet with conservative and liberal constituents separately or together. A CEO usually faces a dilemma about how to announce the firm's profitability given the possible labor negotiations on one side and the bond ratings on the other. Another example would be an entrepreneur trying to convince two investors, who might have different interests in the project, to fund his project. Should the entrepreneur pay separate visits to the two investors or would making promises in the presence of both increase his credibility? The examples are numerous, but they all illustrate the same point that there must be a difference between whispering to people's ears and shouting out loud on the street. The most important difference is regarding how much information is revealed through these two scenarios. One might naturally expect the sender to tailor his announcement depending on the range of interest groups he addresses.

In this paper we present a model to analyze the impact of private and public announcements on information revelation. We take a simplified approach in two ways: (1) we assume that there are two receivers with exactly opposite interests in the information. (2) we restrict the announcements to be verifiable, i.e. the sender is not allowed to lie but can claim that he has no information. In this framework, we first investigate the informativeness of the equilibrium when the seller commits to a private or public announcement. We find that the amount of information conveyed in each case depends on the preferences of the receivers. In particular, our results show that public protocol is more informative if each receiver reacts more aggressively to a possibility of bad news than to an identical possibility of good news for her and the private protocol is more informative if the opposite is true.

It is true that telling something in private and in public has different consequences, but this is not the only thing that matters. It is the sender who decides whether to talk in private or in public. The important question is when he decides. In other words, it also matters whether the seller can commit to a particular protocol of announcement. For some kind of announcements there has developed a convention which restricts the announcement to be in a particular protocol. For example, right after the election the new acclaimed president addresses the whole nation. This is not a matter of choice for the politician. At the other extreme is job interviews. The candidate who is trying to convey information about himself to the employers has a private meeting with each employer. Alternatively you could imagine a career fair where all the employers are watching and each candidate goes on the stage and gives all the necessary information about himself. From a welfare point of view, we ask the question whom these conventions serve. In particular through our model we investigate which commitments the sender and the receivers prefer.

The conventions (commitments) discussed above constitute a very restrictive subset of the information transmission cases. In other words the sender can hardly commit but rather he chooses the protocol after he gets informed. For example, the politician is usually free to have separate meetings with conservatives and liberals. Similarly the CEO can unexpectedly invite the shareholders, without the union representatives being present, for a meeting the day after he is informed about the progress of a project in Middle East. If this is the case though, the receivers, who we assume use all signals to make their inference, will try to see the motive the sender had in choosing the protocol through which they are addressed. We extend our model to allow for no commitment and investigate how much information is revealed in equilibrium. In this case we find that there might be multiple equilibria. However, there is always an equilibrium identical to the equilibrium where the sender commits to a public announcement. This shows that having a public announcement option is almost as effective as committing to it in determining the outcome.

No commitment in fact involves more than what we just described. It is very restrictive to think that the sender's choice is between a one-shot public claim and one-shot private claims. Going back to the politician example, suppose the politician is more informed about consequences of some policy he is going to implement if he gets elected. After he gets informed, up until the voting, he can have as many private talks as he wants with different groups of constituents. He can also address the whole nation on TV. In other words he keeps giving speeches in private and in public till the last minute. This applies to all our previous examples. Therefore, taking one more step towards making better predictions, we also investigate the information transmission when the protocols are unlimited. In this case we find that the power of a public announcement option is eroded by the possibility that the sender can always hold one last private meeting before time ends. Our results show that if equilibrium exists, it implies the same information revelation as the case with commitment to private meetings.

This paper is closely related to Farrell and Gibbons (1989), which develops a model of cheap talk with two audiences to investigate what we referred to as the commitment case. They use a discrete model with two states of the world and two actions available to each receiver. They do not fully characterize equilibria but instead find a stylized relation between the existence of certain type of equilibria in private protocols and in public protocol. The possible scenarios that result from their model could be summarized in three cases: The informativeness of equilibria could be the same in private protocols and public protocol, the informativeness of the equilibria could improve for at least one receiver in the public protocol compared to private protocol, the informativeness of equilibria could deteriorate for one and at most one receiver in the public protocol compared to the private protocol. So their findings allow for improvement with both receivers but not deterioration with both receivers. We use an informatively richer model which at the first glimpse looks like a retreatment of the case in Farrell and Gibbons (1989) where the informativeness is improved for both receivers in public. However, our model allows us to characterize all the equilibria and we show that both mutual improvement and deterioration is possible depending on the characteristics of the receivers. Our model also allows us to address questions that cannot be answered using the Farrell and Gibbons (1989) framework. In particular we are able to extend the model to incorporate the no commitment case and even no commitment with unlimited protocols.

When analyzing the formerly mentioned questions we diverge from Farrell and Gibbons (1989) in terms of the informational assumptions. Farrell and Gibbons (1989) uses a "cheap talk" framework, as introduced by Crawford and Sobel (1982), to analyze this information transmission. In other words, in their model, the sender is allowed to lie and report whatever he wants after getting informed. Instead we follow Shin (2002) and Milgrom and Roberts (1986) in using a "verifiable reports" assumption which restricts the capability of the sender to lie completely. In other words our assumptions imply that ex-post, what the seller has reported should not contradict with the true state of the world. We think this is not too restrictive in the sense that most of the real world examples actually fit better to this case than pure cheap-talk. For a politician, lying about something that soon will be public has irreversible adverse effects on his future political career. Instead politicians prefer to talk more vaguely or to withhold information from the public. Similarly anything the CEO promises should be backed up by some financial document. So we believe that not only does our restriction bring us closer to the real world for most cases, but also it mitigates the problem of multiple equilibria and makes things more tractable.

In Section 2.2, we will describe the model. In Section 2.3, we will characterize the pure strategy equilibria when the sender has committed to private or public announcement ex-ante. In Section 2.4, we will analyze the no commitment issue and finally in Section 2.5, we will extend the analysis to unlimited protocols.

2.2 The Model

2.2.1 Payoff Environment

There is one sender and two receivers. There is a piece of information θ, which can take two possible values $\{-\Theta, \Theta\}$ where $\Theta \in \Re^+$ is a scalar parameter. Receiver i's utility is denoted by $u_r(z_i, \theta_i)$ where $z_i \in \Re$ is the choice variable for receiver $i \in I = \{1, 2\}$. $\theta_i = \alpha_i \theta$ where $\alpha_1, \alpha_2 \in \Re$ are variables that measure how nearly the receivers' interests coincide. The sender's utility is given by $u_s(z_s)$ where $z_s = z_1 + z_2$.

Throughout the model we will assume that the function $u(z_i, \theta_i)$ is continuous, strictly concave in z, and is supermodular with respect to z_i and θ_i. Supermodularity ensures that the receivers' optimal choices increase in θ_i when they are fully informed. We also assume that $-\alpha_1 = \alpha_2 = \alpha$ where $\alpha \in \Re^+$ is a scalar parameter. This assumption implies that the receivers have the exact opposite interests in the information, i.e.

$$\arg\max_{z_1} u_r\left(z_1, -\alpha\left(-\theta\right)\right) = \arg\max_{z_2} u_r\left(z_2, \alpha\theta\right) \tag{2.1}$$

We impose this assumption to be able to highlight the effects of introducing the public announcement into the picture. By doing this we are preparing the most suitable basis for an increased credibility in public compared to private announcements.

On the sender's part we will assume that $u_s'\left(z_s\right) > 0$. This is a monotonicity assumption. Although it might be restrictive for some cases, we believe the sender's interest in the receivers' choices could be represented like this for most of the examples we discussed in the introduction.

2.2.2 Information Structure

Initially both the receivers and sender share the prior

$$\Pr\left(\theta = \Theta\right) = p \tag{2.2}$$

However, the sender observes a signal $\omega\left(\theta\right) \in \{\theta, \emptyset\}$. The conditional probability of each signal is given by

$$\begin{aligned} \Pr\left(\omega = \theta \mid \theta\right) &= q \\ \Pr\left(\omega = \emptyset \mid \theta\right) &= 1 - q \end{aligned} \tag{2.3}$$

We interpret $\omega = \emptyset$ as the sender not observing anything. This information structure implies that the sender either observes a perfect signal with probability q or does not observe anything. The receivers observe neither θ nor ω.

2.2.3 Strategies and Equilibrium Concept

After observing the signal the sender makes an announcement. We are going to analyze this announcement problem in two different ways: a protocol where the sender can interact with each receiver in private and a protocol where the sender is forced to make his announcement in the presence of both receivers, i.e. in public. We first consider the case where the sender has committed to a particular protocol before receiving the information and then we investigate the case where his protocol decision is endogenous, i.e. he decides after observing the information. We will employ the pure-strategy, perfect Bayesian equilibria as the solution concept. In each case we are going to impose a verifiable reports assumption which will imply that the sender is restricted to announce either what he observed or that he has not observed anything. First let us describe the strategies in each case by the following table. Let $\Omega = \{-\Theta, \Theta, \emptyset\}$ and $\Gamma = \{R, U\}$ where R denotes private and U denotes public.

		Strategies	
		sender's announcement	receiver i's action
commitment	private	$a_i : \Omega \to \Omega$ for each $i \in \{1,2\}$	$z_i : \Omega \to \mathbf{R}$
	public	$a : \Omega \to \Omega$	$z_i : \Omega \to \mathbf{R}$
no commitment		$a_i : \Omega \to \Gamma \times \Omega$ for each $i \in \{1,2\}$	$z_i : \Gamma \times \Omega \to \mathbf{R}$

Table 2.1: The Strategies

Notice that a_i and a denote the announcement of the sender once he receives the signal $\omega \in \Omega$. When the protocol is predetermined as private or public, the sender only chooses what to announce. When there is no commitment the sender chooses both the protocol and the reports. We impose a restriction on the sender's strategies: verifiability of the reports, which implies that the sender cannot misclaim any information , i.e. he can only withhold information by claiming that he has not received it. Two other restrictions arise naturally for the no commitment case: (i) if the sender chooses the public protocol, his announcement has to be uniform across receivers, (ii) each receiver has to receive an announcement. These restrictions on the strategies are formalized through the following tables.

		Verifiability Assumption (VA)
commitment	private	$a_i(\omega) \in \{\omega, \emptyset\}$ for all $i \in \{1,2\}$ and $\omega \in \Omega$
	public	$a(\omega) \in \{\omega, \emptyset\}$ for all $\omega \in \Omega$
no commitment		$a_i(\omega) \in \{(\gamma, \widehat{\omega}) \mid \gamma \in \Gamma \text{ and } \widehat{\omega} \in \{\omega, \emptyset\}\}$ for all $\omega \in \Omega, i \in I$

Table 2.2: Verifiability Assumption

	One Announcement Requirement (OA)	**Uniform Public Announcement (UP)**
no commitment	For all $\omega \in \Omega$, $\gamma \in \Gamma$ and $i \in I$, if $a_i(\omega) = (., \gamma)$ then $a_{i'}(\omega) = (., \gamma)$	For all $\omega \in \Omega$ and $i \in I$, if $a_i(\omega) = (U, .)$ then $a_i(\omega) = a_{i'}(\omega)$

Table 2.3: The Restrictions for No Commitment Case

Each receiver chooses an action, z_i, given what she observes, a_i (or a, but from now on we will use a_i for both and the public case will imply $a_1 = a_2$). Notice that in the commitment case the receiver only hears an announcement, whereas in the no commitment case she also observes the protocol choice. Having observed a_i, receiver i computes a posterior for the true state of the world using Bayesian updating, i.e.

$$\widehat{p}_i(a_i) = \Pr(\theta = \Theta \mid a_i) \tag{2.4}$$

Then she chooses an action to maximize her expected payoff. For receiver 2 the optimal action, when she holds a posterior $\widehat{p}$ is given by

$$\xi(\widehat{p}) = \arg\max_z [(1-\widehat{p})\, u(z, -\alpha\Theta) + \widehat{p} u(z, \alpha\Theta)] \tag{2.5}$$

Since we assumed that $\alpha_1 = -\alpha$, we can derive receiver 1's optimal action for the same belief as $\xi(1-\widehat{p})$. From now on we will identify the receivers by their choices. Let us introduce the notation

$$\begin{aligned} \bar{i}(\theta) &\equiv \left\{ i \mid \arg\max_z u(z, \alpha_i\theta) = z(1),\ i \in I \right\} \\ \underline{i}(\theta) &\equiv \left\{ i \mid \arg\max_z u(z, \alpha_i\theta) = z(0),\ i \in I \right\} \end{aligned} \tag{2.6}$$

for $\theta \in \{-\Theta, \Theta\}$. In other words for a given θ, $\bar{i}$ is the receiver who likes the information and chooses the highest action, whereas $\underline{i}$ is the receiver who does not like it and hence chooses the lowest action.

In this setting a pure strategy perfect Bayesian equilibrium is defined by an announcement rule for the sender, $(a_1^*(\omega), a_2^*(\omega))$, an action choice and a belief system for the receivers, $z_i^*(a_i)$ and $\widehat{p}_i^*(a_i)$. These have to satisfy the assumptions and restrictions given above. Notice also that $a_i^*(.)$ is defined differently depending on which case in Table 1 is being analyzed. They also have to satisfy optimality conditions, i.e.:

1. *Optimality of sender's choice:* For all $\omega \in \Omega$

$$\begin{aligned}\left(a_1^*(\omega), a_2^*(\omega)\right) &= \sup_{a_i} u_s\left(z_1^*(a_1)\right) + z_2\left(z_2^*(a_2)\right) \\ &\quad \text{subject to (VA), (OA) and (UP)}\end{aligned} \tag{2.7}$$

2. *Optimality of receiver i's choice*: For all a_1, a_2,

$$\begin{aligned} z_2^*(a_2) &= \xi\left(\widehat{p}_2^*(a_2)\right) \\ z_1^*(a_1) &= \xi\left(1-\widehat{p}_1^*(a_1)\right)\end{aligned} \tag{2.8}$$

   where $\xi(.)$ is as defined in equation (2.5)

3. *Consistency of beliefs:* The beliefs are derived by Bayes Rule on equilibrium path, i.e.

$$\widehat{p}_i^*(a_i) = \frac{\Pr\left(\omega \in \Omega \mid a_i^*(\omega) = a_i, \theta = \Theta\right)p}{\Pr\left(\omega \in \Omega \mid a_i^*(\omega) = a_i, \theta = \Theta\right)p + \Pr\left(\omega \in \Omega \mid a_i^*(\omega) = a_i, \theta = -\Theta\right)(1-p)} \tag{2.9}$$

   for all a_i in the support of $a_i^*(\omega)$. Regarding the off-equilibrium beliefs we assume that $\widehat{p}_1^*(a_1) = \widehat{p}_2^*(a_2)$ if $a_1 = a_2$, i.e. the two receivers have the same off-equilibrium beliefs when they observe the same thing.

2.3 Commitment Equilibria

As we described earlier, in some cases we see either a convention or rule that restricts the sender to one particular protocol. Using our model we will try to characterize the equilibria when the seller is committed to private and public protocols. This is the case analyzed by Farrell and Gibbons (1989).

2.3.1 Private Protocol

Here we seek to find how much information is revealed when the sender is restricted to make private meetings with the two receivers. An example is a politician having committed to make separate announcements to two different groups of constituents about a resource allocation plan. The following proposition reveals the unique equilibrium in this case.

Proposition 2.1. *The pure-strategy perfect Bayesian equilibrium is unique and characterized by*

(i) $a^*_{\bar{i}(\theta)}(\theta) = \theta$, $a^*_{\underline{i}(\theta)}(\theta) = \emptyset$

(ii) $z^*_{\bar{i}(\theta)}(\theta) = z(1)$, $z^*_{\underline{i}(\theta)}(\theta) = z(0)$, $z^*_1(\emptyset) = \xi(1-\widehat{p}_1)$ *and* $z^*_2(\emptyset) = \xi(\widehat{p}_2)$ *for all* $\theta \in \{-\Theta, \Theta\}$ *and* $i \in I$, *where*

$$\widehat{p}_1 = \Pr(\theta = \Theta \mid a_1 = \emptyset) = \frac{p}{(1-q)(1-p)+p} > p \tag{2.10}$$

$$\widehat{p}_2 = \Pr(\theta = \Theta \mid a_2 = \emptyset) = \frac{(1-q)p}{(1-q)p+1-p} < p \tag{2.11}$$

Proof. Suppose there exists a $\theta \in \{-\Theta, \Theta\}$ and an equilibrium where $a^*_{\bar{i}(\theta)}(\theta) = \emptyset$. Upon hearing this the belief of receiver $\bar{i}(\theta)$ will be $\Pr(\theta = \theta \mid \emptyset) < 1$. But then the $z^*_{\bar{i}(\theta)}(\emptyset) < z(1)$ because $\frac{\partial^2 u(z,\theta)}{\partial z \partial \theta} > 0$. Then the sender would rather deviate and report θ. Thus in all equilibria it has to be the case that $a^*_{\bar{i}(\theta)}(\theta) = \theta$. Suppose there exists an equilibrium and a $\theta \in \{-\Theta, \Theta\}$ where $a^*_{\underline{i}(\theta)}(\theta) = \theta$. Then, $z^*_{\underline{i}(\theta)}(\theta) = z(0)$. Suppose the sender deviates to

$a^*_{\underline{i}(\theta)}(\theta) = \emptyset$. Upon hearing this the belief of receiver $\underline{i}(\theta)$ will be $\Pr(\theta = \theta \mid \emptyset) < 1$ and thus $z^*_{\underline{i}(\theta)}(\emptyset) > z(0)$. So the sender would deviate. Given these then the unique pure-strategy equilibrium is as given in the proposition. The beliefs are derived by Bayes rule. ∎

$\widehat{p}_i(p, q)$ in equations (2.10) and (2.11) denotes the updated belief of receiver i after the seller claims that he has not observed any information. The proposition implies that the unique equilibrium in private protocol is partial pooling. In equilibrium the sender announces what the receiver wants to hear if he has that information and otherwise withholds the information. Therefore the receivers act pessimistically when the seller does not announce anything. In other words $\xi(0) < \xi(\widehat{p}_2) < \xi(p) < \xi(1)$ and $\xi(0) < \xi(1 - \widehat{p}_1) < \xi(1 - p) < \xi(1)$.

Returning to our politician example, this means the politician when meeting with one group of constituent will talk all about the resources he is planning to put in their preferred uses and not talk about it at all if he is planning to put more resources in the other group's preferred uses. But the model suggests that withholding information from a group of constituents does not mean they will simply think the politician has not made the plan in detail, on the contrary each group will increase its doubt that the politician is planning to serve the other group's interest.

As we said Farrell and Gibbons (1989) do not characterize the equilibria and allow for a range of possibilities in the private protocol. We instead, employed a relatively restricted model and found that the only pure-strategy equilibrium in private protocol is partial-pooling. Now we would like to see how the informativeness of the equilibrium will change when the sender is restricted to a public protocol.

2.3.2 Public Protocol

According to Farrell and Gibbons (1989), the equilibria in private and public protocol are related in the following sense: Separating equilibria with both receivers in private implies separating equilibria in public and everything else that does not violate this is possible. In other words if there is separating equilibrium with one receiver in private and pooling with the other in private, the equilibrium in public could go either way. Because their model is $2 \times 2 \times 2$, they do not have partial pooling cases like we do. Their result does not leave room for a possibility for the communication to go worse in public with both receivers. Our intention is not to cover all possible cases but we would like to show that in our model, depending on some characterization of the receivers' utilities, the communication could become better or worse for both receivers in public protocol. Referring to our politician example, the question we would like to ask is whether more or less information about the politician's plan be revealed if press made all the correspondences and announcements of the politician public. We show that the answer could be positive or negative depending on what the constituents infer from the unsaid.

We first classify the receivers into two groups as prudent and haphazard depending on their preferences as follows:

$$\begin{array}{l} \textit{Lottery's loss assumptions } (L) \\ \textit{(Prudent)} \end{array} : \begin{cases} (1)\ \frac{\partial^3 u(z,.)}{\partial z^3} \geqslant 0 \\ (2)\ \text{If } \widehat{\theta} > \widehat{\theta}',\ \frac{\partial^2 u(z,\widehat{\theta})}{\partial z^2} \geqslant \frac{\partial^2 u(z,\widehat{\theta}')}{\partial z^2} \end{cases} \tag{2.12}$$

$$\begin{array}{l} \textit{Lottery's gain assumptions } (G) \\ \textit{(Haphazard)} \end{array} : \begin{cases} (1)\ \frac{\partial^3 u(z,.)}{\partial z^3} \leqslant 0 \\ (2)\ \text{If } \widehat{\theta} > \widehat{\theta}',\ \frac{\partial^2 u(z,\widehat{\theta})}{\partial z^2} \leqslant \frac{\partial^2 u(z,\widehat{\theta}')}{\partial z^2} \end{cases} \tag{2.13}$$

where in both cases one of the two inequalities is strict. Both (L) and (G) are assumptions on third order conditions. Suppose that the receiver already knows what true θ is. Think about a lottery which says θ will stay at its current value with some probability and change with the remaining probability. (L) implies that the receiver would change his action more aggressively if this is a downward lottery and less aggressively if this is an upward lottery (downward and upward refer to the receivers preference over the true state of the world). That is why we will refer to receivers who satisfy this condition as "prudent". Similarly (G) implies that the receiver would change his action less aggressively if this is a downward lottery and more aggressively if this is an upward lottery. We will refer to receivers who satisfy this condition as "haphazard".

Proposition 2.2. *If the receivers are prudent, the pure-strategy perfect Bayesian equilibrium is unique. It is truthful and is characterized by $a^*(\omega) = \omega$ for all $\omega \in \Omega$.*

Proof. First we will show that it is an equilibrium. Suppose the sender observes $\omega = \theta$ for some $\theta \in \{-\Theta, \Theta\}$. Then under the truthful equilibrium $a^*(\theta) = \theta$. And thus $z^*_{\bar{\imath}(\theta)}(\theta) = \xi(1)$, $z^*_{\underline{i}(\theta)}(\theta) = \xi(0)$. So the sender gets $u_s(\xi(1) + \xi(0))$. Alternatively the sender might want to announce $\emptyset$ after observing θ. We will show that this gives a lower utility to the sender than reporting θ. In the truthful equilibrium when the sender reports $\emptyset$, the receivers choose $z_1^*(\emptyset) = \xi(1-p)$ and $z_2^*(\emptyset) = \xi(p)$.

Suppose $\xi(1) - \xi(p) \leqslant \xi(1-p) - \xi(0)$. We will show that this cannot happen.

There are two cases that we have to go through.

i) Suppose $p \geqslant 1-p$. This means $\xi(p) \geqslant \xi(1-p)$

$$\begin{aligned}
\frac{\partial u(\xi(p),\Theta)}{\partial z} - \frac{\partial u(\xi(1),\Theta)}{\partial z} &= -\int_{\xi(p)}^{\xi(1)} \frac{\partial^2 u(z,\Theta)}{\partial z^2}dz \leqslant -\int_{\xi(p)}^{\xi(1)} \frac{\partial^2 u(z,-\Theta)}{\partial z^2}dz \quad (2.14)\\
&< -\int_{\xi(0)}^{\xi(1-p)} \frac{\partial^2 u(z,-\Theta)}{\partial z^2}dz\\
&= \frac{\partial u(\xi(0),-\Theta)}{\partial z} - \frac{\partial u(\xi(1-p),-\Theta)}{\partial z}
\end{aligned}$$

The first inequality is due to $(L)(1)$ and the second strict inequality is due to $(L)(2)$. But we also know that $\frac{\partial u(\xi(1),\Theta)}{\partial z} = \frac{\partial u(\xi(0),-\Theta)}{\partial z} = 0$, because $\xi(1) = \arg\max_z u(z,\Theta)$ and $\xi(0) = \arg\max_z u(z,-\Theta)$. So the inequality above reduces to

$$\frac{\partial u(\xi(p),\Theta)}{\partial z} < -\frac{\partial u(\xi(1-p),-\Theta)}{\partial z} \quad (2.15)$$

Now, let us look at the change in the following difference:

$$\left[p\frac{\partial u(z,-\Theta)}{\partial z} + (1-p)\frac{\partial u(z,\Theta)}{\partial z}\right] - \frac{\partial u(z,-\Theta)}{\partial z} = (1-p)\left[\frac{\partial u(z,\Theta)}{\partial z} - \frac{\partial u(z,-\Theta)}{\partial z}\right] \quad (2.16)$$

$$\frac{\partial}{\partial z}\left[(1-p)\left[\frac{\partial u(z,\Theta)}{\partial z} - \frac{\partial u(z,-\Theta)}{\partial z}\right]\right] = (1-p)\left[\frac{\partial^2 u(z,\Theta)}{\partial z^2} - \frac{\partial u^2(z,-\Theta)}{\partial z^2}\right] \geqslant 0 \quad (2.17)$$

The last inequality is due to $(L)(1)$.

So the difference between $\left[p\frac{\partial u(z,-\Theta)}{\partial z}+(1-p)\frac{\partial u(z,\Theta)}{\partial z}\right]$ and $\frac{\partial u(z,-\Theta)}{\partial z}$ is weakly increasing in z.

Now let us look at this difference at $\xi(p)$. Since $\xi(p) \geqslant \xi(1-p)$ and the difference is increasing, the difference at $\xi(p)$ will be weakly greater than the difference at $\xi(1-p)$

$$\begin{aligned} &\left[p\frac{\partial u(\xi(p),-\Theta)}{\partial z}+(1-p)\frac{\partial u(\xi(p),\Theta)}{\partial z}\right]-\frac{\partial u(\xi(p),-\Theta)}{\partial z} \\ \geqslant &\left[p\frac{\partial u(\xi(1-p),-\Theta)}{\partial z}+(1-p)\frac{\partial u(\xi(1-p),\Theta)}{\partial z}\right]-\frac{\partial u(\xi(1-p),-\Theta)}{\partial z} \end{aligned} \tag{2.18}$$

But the first term in the brackets on right hand side is zero because,

$\xi(1-p) = \arg\max_z E_\theta[u(z,\theta) \mid \Pr(\theta=\Theta)=1-p]$, so it reduces to

$$\begin{aligned} &\left[p\frac{\partial u(\xi(p),-\Theta)}{\partial z}+(1-p)\frac{\partial u(\xi(p),\Theta)}{\partial z}\right]-\frac{\partial u(\xi(p),-\Theta)}{\partial z} \\ \geqslant &-\frac{\partial u(\xi(1-p),-\Theta)}{\partial z} \end{aligned} \tag{2.19}$$

And from (2.15) we know that $\frac{\partial u(\xi(p),\Theta)}{\partial z} < -\frac{\partial u(\xi(1-p),-\Theta)}{\partial z}$, so

$$\begin{aligned} &\left[p\frac{\partial u(\xi(p),-\Theta)}{\partial z}+(1-p)\frac{\partial u(\xi(p),\Theta)}{\partial z}\right]-\frac{\partial u(\xi(p),-\Theta)}{\partial z} \\ > &\frac{\partial u(\xi(p),\Theta)}{\partial z} \end{aligned} \tag{2.20}$$

Rearranging the terms gives

$$\left[(1-p)\frac{\partial u(\xi(p),-\Theta)}{\partial z}+p\frac{\partial u(\xi(p),\Theta)}{\partial z}\right]<0 \tag{2.21}$$

But the left hand side of this term is equal to 0 because,

$\xi(p) = \arg\max_z E_\theta [u(z,\theta) \mid \Pr(\theta = \Theta) = p]$. So we have $0 < 0$, which is a contradiction. Thus $\xi(1) - \xi(p) > \xi(1-p) - \xi(0)$.

ii) Now suppose $1 - p > p$. This means $\xi(1-p) < \xi(p)$. $\xi(1) - \xi(p) \leqslant \xi(1-p) - \xi(0)$ is equivalent to $\xi(1) - \xi(1-p) \leqslant \xi(p) - \xi(0)$. Now let us suppose $\xi(1) - \xi(1-p) \leqslant \xi(p) - \xi(0)$. The same argument above applies if we switch the places of $\xi(p)$ and $\xi(1-p)$ and of course the places of $1-p$ and p. When we follow that in the end we get

$$\left[p\frac{\partial u(\xi(1-p), -\Theta)}{\partial z} + (1-p)\frac{\partial u(\xi(1-p), \Theta)}{\partial z}\right] < 0 \tag{2.22}$$

But again the left hand side of this term is equal to 0 because it is the maximizer. So, again we get a contradiction $0 < 0$, which means $\xi(1) - \xi(1-p) > \xi(p) - \xi(0)$.

So now, we know that $\xi(1) - \xi(p) > \xi(1-p) - \xi(0)$, which means $\xi(1) + \xi(0) > \xi(p) + \xi(1-p)$. Since $u_s(z_s)$ is increasing in z_s, $u_s(\xi(1) + \xi(0)) > u_s(\xi(p) + \xi(1-p))$. So $a^*(\theta) = \theta$ is indeed an equilibrium.

Uniqueness is very easy to show. We just showed that $\xi(1) + \xi(0) > \xi(p) + \xi(1-p)$ for all $p \in (0,1)$. Any announcement rule for which there exists some $\theta \in \{-\Theta, \Theta\}$ such that $a(\theta) = \emptyset$ would lead to an inference $\widehat{p}(\emptyset) = \Pr(\theta \mid \emptyset) < 1$, so the sum of receivers choice would be $\xi(\widehat{p}) + \xi(1-\widehat{p})$ which as we proved is strictly smaller than $\xi(1) + \xi(0)$ for all $\widehat{p}$. So the sender would rather deviate to $a(\theta) = \theta$. ∎

The proposition says that when the receivers are prudent, i.e. when (L) is true, all information is revealed through the public announcement. The intuition is as follows. If the sender is truthful, one receiver will choose the highest action and the other will choose the lowest action. The sender's other alternative is claiming that he has not observed anything. If he does that, both receivers will keep their prior about the true state of the world. Hence, one receiver will choose a higher action than the truthful report case and the other will choose a lower action. But since the receivers are prudent, the receiver who shades his action down responds more aggressively than the receiver who shades his action up. Formally, using (L) the proof establishes the fact that

$$\xi(1) + \xi(0) > \xi(p) + \xi(1-p) \tag{2.23}$$

for all $p \in (0,1)$. Therefore the total action is smaller when the sender hides information and hence the sender gets a lower payoff.

With prudent receivers we can conclude that having a public announcement increases the credibility of the sender. This resembles the case that Farrell and Gibbons (1989) refers to as "mutual discipline",where the equilibrium is pooling with each receiver in private but separating in public. The informativeness of equilibria increased for both receivers in the public protocol. In other words the existence of each receiver disciplines the communication of the sender with the other receiver.

Proposition 2.3. *If the receivers are haphazard, the pure-strategy perfect Bayesian equilibrium is unique. It is non-revealing (pooling) and is characterized by* $a^*(\omega) = \emptyset$ *for all* $\omega \in \Omega$.

Proof. The fact that the characterization above is an equilibrium is given by following the same proof above with all the inequalities reversed. When all the inequalities are reversed due to (G) assumption, it follows that $\xi(1) + \xi(0) < \xi(p) + \xi(1-p)$ for all $p \in (0,1)$. And thus not revealing anything is in fact an equilibrium.

For uniqueness, suppose there exists an equilibrium where for some $\theta \in \{-\Theta, \Theta\}$ $a(\theta) = \theta$. This report would lead to $\xi(1) + \xi(0)$ for the sender. In this equilibrium the belief that is generated when the sender reports $\emptyset$ is given by $0 < \widehat{p}(\emptyset) = \Pr(\theta \mid \emptyset) < 1$. So the sum of receivers choice upon hearing this would be $\xi(\widehat{p}) + \xi(1-\widehat{p}) > \xi(1) + \xi(0)$. But then the sender would rather deviate to $a(\theta) = \emptyset$. ∎

So when the receivers are haphazard, i.e. when (G) is true, no information is revealed through the public announcement. The intuition is that if the sender acts truthfully, then one receiver chooses the highest action at the expense of the other receiver who chooses the lowest action. Alternatively the sender can withhold information, in which case the receivers will hold onto their initial prior and hence one receiver's action will increase whereas the other's will decrease. Since the receivers are haphazard, the receiver shading his action up will do so more aggressively than the receiver shading his value down. Formally, using (G) the proof establishes the fact that

$$\xi(1) + \xi(0) < \xi(p) + \xi(1-p) \tag{2.24}$$

for all $p \in (0,1)$. Hence, the total action increases when the sender hides information and the sender is better off.

This is a case that could not be discussed by Farrell and Gibbons because their model allowed for only pure pooling or separating equilibria. In the previous section, we showed that when the announcements were made separately to the two receivers, the equilibria was partially-pooling, so at least some information was revealed to each receiver. But now, the informativeness of the equilibria decreases for both receivers in the public protocol.

Proposition 2.3 reveals that even when the receivers have exactly opposite interests on the information that the sender has, the credibility of the sender is not necessarily increased thorough a public announcement.

Returning to our politician example, if the press had access to all the reports of the politician, how much the politician would reveal about his plans would depend on his perception of the constituents. In other words if the politician thought the constituents are prudent in their inference, he would reveal plans that he would not reveal in private interviews. However, if the politician thought the constituents are haphazard, he would always keep silent.

Notice that we investigated the two extreme cases, namely the cases that satisfy (L) and (G) for both receivers. There are numerous other cases that satisfy neither, i.e. that are in between. However, our intention is not to cover all possibilities but to show that different perceptions about the receivers by the sender might lead to completely different outcomes.

2.3.3 Welfare analysis

The question we would like to ask here is which protocol is preferred by the sender and the receivers. Receivers clearly prefer the more informative equilibrium. In other words, the prudent receivers prefer the public protocol whereas haphazard receivers prefer the private protocol. It is more complicated to answer the same question from the sender's point of view. We first fully characterize the ex-post preference for the sender. Then we discuss his ex-ante preferences. For haphazard receivers we are able to characterize the ex-ante preferences of the sender for cases where there is a high probability that the sender gets informed. We cannot determine the preferences for low probability levels. For prudent receivers, all we can determine is that as the receivers get more confident about the informational status of the sender, the sender's expected payoff from private and public protocols converge.

Lemma 2.1. *If the receivers are prudent, the sender ex-post prefers*

(*i*) *the public protocol when $\omega = \emptyset$, i.e. whenever he does not get any information and*

(*ii*) *the private protocol whenever $\omega \in \{\Theta, -\Theta\}$, i.e. whenever he gets information.*

Proof. (*i*) Suppose the sender observes $\omega = \emptyset$. With the private protocol $z_1^*(\emptyset) = \xi(1 - \widehat{p}_1)$ and $z_2^*(\emptyset) = \xi(\widehat{p}_2)$, where $\widehat{p}_1$ and $\widehat{p}_2$ are as defined in Proposition 1. So the sender's utility is $u_s(\xi(1-\widehat{p}_1) + \xi(\widehat{p}_2))$. With the public protocol, in the truthful equilibrium $z_1^*(\emptyset) = \xi(1-p)$ and $z_2^*(\emptyset) = \xi(p)$, thus the sender's utility is $u_s(\xi(1-p) + \xi(p))$. Since $\xi(1-p) > \xi(1-\widehat{p}_1)$ and $\xi(p) > \xi(1-\widehat{p}_1)$, $u_s(\xi(1-p)+\xi(p)) > u_s(\xi(1-\widehat{p}_1) + \xi(\widehat{p}_2))$.

(*ii*) Suppose the seller observes $\omega = \theta \in \{\Theta, -\Theta\}$. With the private protocol $z^*_{\bar{i}(\theta)}(\theta) = \xi(1)$ and $z^*_{\underline{i}(\theta)}(\emptyset) = \xi\left(p_{\underline{i}(\theta)}\right)$ where $p_{\underline{i}(\Theta)} = 1 - \widehat{p}_1$ and $p_{\underline{i}(-\Theta)} = \widehat{p}_2$, where $\widehat{p}_1$ and $\widehat{p}_2$ are as given before. So sender's utility with the private protocol is $u_s\left(\xi(1) + \xi\left(p_{\underline{i}(\theta)}\right)\right)$. With the public protocol, in the truthful equilibrium $z^*_{\bar{i}(\theta)}(\theta) = \xi(1)$, $z^*_{\underline{i}(\theta)}(\theta) = \xi(0)$. So the sender's utility with the public protocol will be $u_s(\xi(1) + \xi(0))$. Since $\xi\left(p_{\underline{i}(\theta)}\right) > \xi(0)$, $u_s\left(\xi(1) + \xi\left(p_{\underline{i}(\theta)}\right)\right) > u_s(\xi(1) + \xi(0))$, thus the sender prefers the private protocol whenever he gets information. ∎

In other words, the sender prefers the public protocol in the cases where he does not get any information. The intuitive reason for this is that, not announcing anything leads to a more pessimistic inference by the receivers in the private protocols compared to the inference in public protocol. Similarly the sender prefers a private protocol when he observes information, because the public announcement results in fully sacrificing one receiver whereas the same receiver responds more optimistically when she is not told anything in the private protocol.

Lemma 2.2. *If the receivers are haphazard,*

(*i*) *the sender ex-post prefers the public protocol when* $\omega = \emptyset$

(*ii*) *there exists* $\underline{q}(p) < \overline{q}(p) \in (0,1)$ *such that when* $\omega \in \{\Theta, -\Theta\}$ *the sender ex-post prefers*

(1) *the private protocol for all* $q \leqslant \underline{q}(p)$

(2) *the public protocol for all* $q \geqslant \overline{q}(p)$.

(3) *the private protocol if* $\Pr(\theta = \omega) \leqslant \frac{1}{2}$ *and public otherwise, for all* $q \in \left(\underline{q}(p), \overline{q}(p)\right)$

Proof. (*i*)Suppose the sender observes $\omega = \emptyset$. With the private protocol $z_1^*(\emptyset) = \xi(1 - \widehat{p}_1, \widehat{p}_1)$ and $z_2^*(\emptyset) = \xi(\widehat{p}_2, 1 - \widehat{p}_2)$. So the sender's utility is $u_s\left(\widehat{\xi}(1 - \widehat{p}_1) + \widehat{\xi}(\widehat{p}_2)\right)$. With the public protocol, in the non-revealing equilibrium, $z_1^*(\emptyset) = \xi(1-p)$ and $z_2^*(\emptyset) = \xi(p)$, thus the sender's utility is $u_s(\xi(1-p) + \xi(p))$. Since $\xi(1-p) > \xi(1 - \widehat{p}_1)$ and $\xi(p) > \xi(\widehat{p}_2)$, $u_s(\xi(1-p) + \xi(p)) > u_s(\xi(1 - \widehat{p}_1) + \xi(\widehat{p}_2))$.

(*ii*) Let $v(\pi) = \xi(1) + \xi(\pi)$ for $\pi \in [0,1]$. Notice that $v(\pi)$ is increasing in π. Since $\xi(1) > \xi(\pi)$ for any $\pi \in (0,1)$, $\xi(1) + \min[\xi(1-p), \xi(p)] > \xi(p) + \xi(1-p)$. By equation (2.24) $\xi(1) + \xi(0) < \xi(1-p) + \xi(p)$. Thus $v(0) < \xi(1-p) + \xi(p) < v(p)$. Since $u_r(z,.)$ is continuous in z, and $z \in \mathbf{R}$, $\xi(\pi)$ is continuous in π. By intermediate value theorem there has to be a $0 < \pi^*(p) < \min(1-p, p)$ such that $v(\pi^*(p)) = \xi(1) + \xi(\pi^*(p)) = \xi(1-p) + \xi(p)$.

We also know that $\widehat{p}_i(p,q)$ as given in Proposition 2.1, is continuous in q for $i \in I$ with $1 - \widehat{p}_1(p,0) = 1 - p$, $\widehat{p}_2(p,0) = p$ and $1 - \widehat{p}_1(p,1) = \widehat{p}_2(p,1) = 0$. So again by intermediate value theorem there is a $0 < q_1(p) < 1$ and a $0 < q_2(p) < 1$ such that $1 - \widehat{p}_1(p, q_1(p)) = \widehat{p}_2(p, q_2(p)) = \pi^*(p)$. Since both $1 - \widehat{p}_1(p,q)$ and $\widehat{p}_2(p,q)$ are decreasing in q, $1 - \widehat{p}_1(p,q) \lesseqgtr \pi^*(p)$ and $\widehat{p}_2(p,q) \lesseqgtr \pi^*(p)$ as $q \gtreqless \overline{q}(p)$ and $q \gtreqless \overline{q}(p)$ respectively. Since $v(\pi)$ is increasing in π, $v(1 - \widehat{p}_1(p,q)) = \xi(1) + \xi(1 - \widehat{p}_1(p,q)) \lesseqgtr \xi(1-p) + \xi(p)$ and $v(\widehat{p}_2(p,q)) = \xi(1) + \xi(\widehat{p}_2(p,q)) \lesseqgtr \xi(1-p) + \xi(p)$ as $q \gtreqless q_1(p)$ and $q \gtreqless q_2(p)$ respectively. Thus when $\theta = -\Theta$, the sender ex-post prefers private protocol if $q < q_2(p)$ and public protocol if $q > q_2(p)$. Similarly when $\theta = \Theta$, the sender ex-post prefers private protocol if $q < q_1(p)$ and public protocol if $q > q_1(p)$.

In other words at the interim stage, when the sender learns that $\omega \in \{\Theta, -\Theta\}$, if $q > \max\{q_1(p), q_2(p)\} = \overline{q}(p)$ he prefers the public protocol and if $q < \min\{q_1(p), q_2(p)\} = \underline{q}(p)$ he prefers the private protocol. If $\underline{q}(p) < q < \overline{q}(p)$, he prefers different protocols upon receiving $\omega = -\Theta$ and $\omega' = \Theta$. If $p > \frac{1}{2}$, it will be the case that $\widehat{p}_2(p,q) > 1 - \widehat{p}_1(p,q)$ for all p and $q \in (0,1)$. Since both $\widehat{p}_2(p,q)$ and $1 - \widehat{p}_1(p,q)$ are decreasing in q, $q_1(p) < q_2(p)$. Thus when $q_1(p) < q < q_2(p)$, the sender ex-post prefers the private protocol if $\omega = -\Theta$ and the public protocol if $\omega = \Theta$. If on the other hand $p \leqslant \frac{1}{2}$, when $q_2(p) < q < q_1(p)$, the sender will ex-post prefer the public protocol if $\omega = -\Theta$ and the private protocol if $\omega = \Theta$. ∎

The sender always prefers the public protocol ex-post when he does not get any information. However his ex-post preference when he is informed depends on what kind of receivers he is facing. In particular, with prudent receivers he is always better off with private protocol when he gets informed, but with haphazard receivers he might prefer public protocol at least for one type of signal if the receivers have sufficient belief in the fact that the sender is informed.

Having discussed the ex-post preference, we will analyze the ex-ante preferences of the sender. Sender's expected payoff in the equilibrium with both haphazard and prudent receivers in private protocol is

$$\begin{aligned} Eu_s^R(q,p) &= (1-q)\,u_s\left(\xi\left(1-\widehat{p}_1\right)+\xi\left(\widehat{p}_2\right)\right)+q\left(1-p\right)u_s\left(\xi\left(1\right)+\xi\left(\widehat{p}_2\right)\right) \\ &\quad +qpu_s\left(\xi\left(1-\widehat{p}_1\right)+\xi\left(1\right)\right) \end{aligned} \tag{2.25}$$

where $\widehat{p}_1$ and $\widehat{p}_2$ are as defined in Proposition 2.1.

Sender's expected payoff in the equilibrium with prudent receivers in public protocol is

$$Eu_s^{U,P}(q,p) = (1-q)\,u_s\left(\xi\left(1-p\right)+\xi\left(p\right)\right)+qu_s\left(\xi\left(1\right)+\xi\left(0\right)\right) \tag{2.26}$$

Sender's expected payoff in the equilibrium with haphazard receivers in public protocol is

$$Eu_s^{U,H}(q,p) = u_s\left(\xi\left(1-p\right)+\xi\left(p\right)\right) \tag{2.27}$$

Corollary 2.1. *If the receivers are haphazard and if $q > \overline{q}(p)$ the sender ex-ante prefers the public protocol.*

Proof. The proof follows from Lemma 2.2. If the receivers are haphazard and $q > \overline{q}(p)$, the sender prefers the public protocol ex-post for all $\omega \in \Omega$ and thus he prefers it ex-ante. ∎

Thus, if the receivers are haphazard and there is a good chance that the sender will be more informed than the receivers, the sender had better commit to a public announcement.

The next proposition reveals a property of the ex-ante preference when the receivers are prudent.

Proposition 2.4. $\lim_{q\to 0}\left(Eu_s^R(q,p) - Eu_s^{U,P}(q,p)\right) = \lim_{q\to 1}\left(Eu_s^R(q,p) - Eu_s^{U,P}(q,p)\right)$ *and they both equal zero.*

Proof. $\lim_{q\to 0}\widehat{p}_1(p,q) = \lim_{q\to 0}\widehat{p}_2(p,q) = p$ and $\lim_{q\to 1}(1-\widehat{p}_1(p,q)) = \lim_{q\to 1}\widehat{p}_2(p,q) = 0$. Since $\xi(.)$ is continuous, $\lim_{q\to 0}\xi(1-\widehat{p}_1(p,q)) = \xi(1-p)$, $\lim_{q\to 0}\xi(\widehat{p}_2(p,q)) = \xi(p)$ and $\lim_{q\to 1}\xi(1-\widehat{p}_1(p,q)) = \lim_{q\to 1}\xi(\widehat{p}_2(p,q)) = \xi(0)$. Continuity of the argument of the limit gives us the result. ∎

As it gets more and more probable that the sender will not observe anything, in the private protocol the prudent receivers tend to believe it more when the sender says he has not observed anything. Thus the outcome of the private protocol and the public protocol converge. Similarly when it is almost certain that the sender will get information, this time the receiver who gets to hear nothing in private makes a choice very close to what he would choose in the most prudent scenario. Thus again the outcome of the private protocol and private protocol converge.

2.4 No Commitment Equilibria

As we argued in the introduction there are some occasions where the sender does not have the chance to commit. For example, apart from a few announcements per year the politician has to make in public, there is no restriction on the type of the announcement he makes. He gets informed by one of his advisors about a particular project and then he plans how to convey this to the public. This is the question that is left unanswered by Farrell and Gibbons (1989). In this section of the paper we intend to analyze the information transmission when the receivers know that the sender cannot commit and we will try to see how the informativeness of equilibria is affected by no commitment. In particular as shown in Table 2.1 the sender chooses the protocol at the interim stage, i.e. after he gets his signal. Hence he simultaneously chooses the protocol and the announcement. The receivers are aware of that. Hence, as shown in Table 2.1, they will try to infer information from the protocol choice as well. In other words now the a_i in equation (2.9) is two dimensional, i.e. it incorporates both the protocol and the announcement. Similar to the commitment case, we investigate the problem from two different angles, namely prudent receivers and haphazard receivers.

2.4.1 Equilibria with prudent receivers

In this section we investigate information transmission to prudent receivers when the sender cannot commit to a protocol. We first show that in equilibrium the sender always prefers to reveal some information to not revealing at all. Then we show that in equilibrium it can not be the case that one receiver gets more informed than the other regardless of the signal. Using these results we show that there are two possible equilibria, one where each receiver gets informed about only one signal and another one where they get perfectly informed about both signals. We find that the first kind of equilibrium exists for a restricted set of parameters whereas the second one is always an equilibrium.

Lemma 2.3. *There exists no pure-strategy perfect Bayesian equilibrium that is non-revealing (pooling) with prudent receivers, i.e. there exists no equilibrium where for all* $\theta \in \{\Theta, -\Theta\}$, *for all* $i \in I$ *and for some* $\gamma \in \Gamma$, $a_i^*(\theta) = (\gamma, \emptyset)$.

Proof. Suppose the sender observes $\omega = \theta \in \{\Theta, -\Theta\}$. In equilibrium $u_s(\xi(1-p)+\xi(p))$ is what he gets. Suppose he deviates to $a_{\bar{i}(\theta)}(\theta) = (R,\theta)$, $a_{\underline{i}(\theta)}(\theta) = (R,\emptyset)$. The worst he could get in this case, depending on the off-equilibrium beliefs is $u_s(\xi(1)+\xi(0)) > u_s(\xi(1-p)+\xi(p))$, since $\xi(1)+\xi(0) > \xi(1-p)+\xi(p)$ by equation (2.23). So he would rather deviate. ∎

Lemma 2.3 establishes the fact that when the receivers are prudent, the sender will always prefer to give information to at least one receiver through a private protocol. Hence in equilibrium some information is transmitted.

Lemma 2.4. *There exists no pure-strategy perfect Bayesian equilibrium with prudent receivers where one receiver in equilibrium gets more information than the other, i.e. in all equilibria if there exists $\theta, \theta' \in \{\Theta, -\Theta\}$ and $\omega \in \{\theta', \emptyset\}$ such that $a^*_{\bar{i}(\theta)}(\theta) = (R,\theta)$, $a^*_{\underline{i}(\theta)}(\theta) = a^*_{\underline{i}(\theta)}(\omega) = (R,\emptyset)$ then $a^*_{\underline{i}(\theta)}(\theta') = (\gamma, \theta')$ for some $\gamma \in \Gamma$ and $a^*_{\bar{i}(\theta)}(\theta') = a^*_{\bar{i}(\theta)}(\emptyset)$.*[1]

Proof. We will start by proving the first clause in the conditional statement. Suppose there exists $\theta, \theta' \in \{\Theta, -\Theta\}$ and $\omega \in \{\theta', \emptyset\}$ such that $a^*_{\bar{i}(\theta)}(\theta) = (R,\theta)$, $a^*_{\underline{i}(\theta)}(\theta) = a^*_{\underline{i}(\theta)}(\omega) = (R,\emptyset)$, but $a^*_{\underline{i}(\theta)}(\theta') = (\gamma,\emptyset)$ for some $\gamma \in \Gamma$. This implies either $\gamma = R$ or $a^*_{\underline{i}(\theta)}(\emptyset) = a^*_{\bar{i}(\theta)}(\emptyset) = (R,\emptyset)$. If $\gamma = R$ then the sender would rather deviate to $a^*_{\underline{i}(\theta)}(\theta') = (R,\theta')$. If $\gamma = U$ and $a^*_{\underline{i}(\theta)}(\emptyset) = a^*_{\bar{i}(\theta)}(\emptyset) = (R,\emptyset)$, then $z_{\bar{i}(\theta)}(U,\emptyset) = \xi(0)$ because receiver $\bar{i}(\theta)$'s equilibrium belief given $a^*_{\bar{i}(\theta)}(\theta) = (R,\theta)$ and $a^*_{\bar{i}(\theta)}(\emptyset) = (R,\emptyset)$ is $\Pr(\theta = \theta' \mid (U,\emptyset)) = 1$. Similarly $z_{\underline{i}(\theta)}(U,\emptyset) = \xi(1)$ because receiver $\underline{i}(\theta)$'s equilibrium belief given $a^*_{\underline{i}(\theta)}(\theta) = (R,\emptyset)$

[1] Notice that the non-existence of such an equilibrium with $\underline{i}(\theta)$ as the more informed receiver in the occurence of θ is very clear. $\underline{i}(\theta)$ being more informed about θ implies $r^*_{\bar{i}(\theta)}(\theta) = (R,\emptyset)$ and thus $z_{\bar{i}(\theta)}(R,\emptyset) < z(\beta)$. Thus the sender would rather deviate to $r^*_{\bar{i}(\theta)}(\theta) = (R,\theta)$.

and $a^*_{\underline{i}(\theta)}(\emptyset) = (R, \emptyset)$ is $\Pr(\theta = \theta' \mid (U, \emptyset)) = 1$. So the sender's utility in equilibrium in this case is $u_s(\xi(-\beta) + \xi(\beta))$. Instead consider the deviation $a_{\underline{i}(\theta)}(\theta') = (R, \theta')$ and $a_{\bar{i}(\theta)}(\theta') = (R, \emptyset)$. In this case, $z_{\underline{i}(\theta)}(R, \theta') = \xi(1)$ and $z_{\bar{i}(\theta)}(R, \emptyset) > \xi(0)$ because $\bar{i}(\theta)$'s equilibrium belief given $a^*_{\bar{i}(\theta)}(\theta) = (R, \theta)$ and $a^*_{\bar{i}(\theta)}(\emptyset) = (R, \emptyset)$ is $\Pr(\theta = \theta' \mid (R, \emptyset)) < 1$. Thus the sender's utility with this deviation $u_s\left(z_{\underline{i}(\theta)}(R, \theta') + z_{\bar{i}(\theta)}(R, \emptyset)\right) > u_s(\xi(0) + \xi(1))$ and the sender would deviate. Hence it has to be the case that $a^*_{\underline{i}(\theta)}(\theta') = (\gamma, \theta')$ for some $\gamma \in \Gamma$.

Having proven the first part, we move to the second statement in the if clause. Suppose there exists $\theta, \theta' \in \{\Theta, -\Theta\}$ and $\omega \in \{\theta', \emptyset\}$ such that $a^*_{\bar{i}(\theta)}(\theta) = (R, \theta)$, $a^*_{\underline{i}(\theta)}(\theta) = a^*_{\underline{i}(\theta)}(\omega) = (R, \emptyset)$. We showed above that this implies $a^*_{\underline{i}(\theta)}(\theta') = (\gamma, \theta')$ for some $\gamma \in \Gamma$. Now suppose $a^*_{\bar{i}(\theta)}(\theta') \neq a^*_{\bar{i}(\theta)}(\emptyset)$. This implies two things: $a^*_{\underline{i}(\theta)}(\emptyset) = a^*_{\bar{i}(\theta)}(\emptyset) = (R, \emptyset)$ and $a^*_{\bar{i}(\theta)}(\theta') = (\gamma, \theta')$ for some $\gamma \in \Gamma$. Notice that $\bar{i}(\theta) = \underline{i}(\theta')$ and $\underline{i}(\theta) = \bar{i}(\theta')$, thus $z_{\bar{i}(\theta)}(\gamma, \theta') = \xi(0)$ and $z_{\underline{i}(\theta)}(\gamma, \theta') = \xi(1)$. The sender in this equilibrium gets $u_s(\xi(0) + \xi(1))$. Consider a deviation to $a^*_{\underline{i}(\theta)}(\theta') = (R, \theta')$ and $a^*_{\bar{i}(\theta)}(\theta') = (R, \emptyset)$. $z_{\underline{i}(\theta)}(R, \theta') = \xi(1)$ and since $a^*_{\bar{i}(\theta)}(\emptyset) = (R, \emptyset)$, receiver $\bar{i}(\theta)$'s belief is $\Pr(\theta = \theta' \mid (R, \emptyset)) < 1$ and thus $z_{\bar{i}(\theta)}(R, \emptyset) > \xi(0)$ and thus $u_s\left(z_{\underline{i}(\theta)}(R, \theta') + z_{\bar{i}(\theta)}(R, \emptyset)\right) > u_s(\xi(0) + \xi(1))$. So the sender would deviate. Hence it has to be the case that $a^*_{\bar{i}(\theta)}(\theta') = a^*_{\bar{i}(\theta)}(\emptyset)$. ∎

The previous two lemmas show that the two receivers have to get at least partially informed and they should get informed in exactly the same number of cases. First let us look at the case where they both get partially informed. Notice that the case where the receivers get partially informed about the same signal is trivial. Suppose they both get informed in private or in public when $\omega = \theta$, but not when $\omega = \theta'$ where θ, $\theta' \in \{\Theta, -\Theta\}$.

Then when $\omega = \theta'$, the receivers' responses will be $\xi(1-\widehat{p}_i) + \xi(\widehat{p}_i)$ for some $i \in I$. By equation (2.23) we know that $\xi(1-\widehat{p}_i) + \xi(\widehat{p}_i) < \xi(0) + \xi(1)$. But then the sender would rather deviate to $a_i(\theta') = (\gamma, \theta')$ for some $\gamma \in \Gamma$ and for all $i \in I$. Having talked about this case, the only possible case left where both receivers get partially informed is where they get partially informed about opposite signals. We will next look for such an equilibrium, which in fact is identical to the equilibrium we had in the case where the sender commits to private protocol.

Lemma 2.5. *There exists a pure-strategy perfect Bayesian equilibrium with prudent receivers where both receivers get partially informed about opposite signals, i.e. an equilibrium where* $a^*_{i(\theta)}(\theta) = (R, \theta)$, $a^*_{\underline{i}(\theta)}(\theta) = (R, \emptyset)$ *and* $a^*_i(\emptyset) = (R, \emptyset)$ *for all* $i \in I$, $\theta \in\in \{\Theta, -\Theta\}$, *if and only if* $2\xi\left(\frac{1}{2}\right) \leqslant \xi(1-\widehat{p}_1) + \xi(\widehat{p}_2)$

Proof. Suppose $2\xi\left(\frac{1}{2}\right) > \xi(1-\widehat{p}_1) + \xi(\widehat{p}_2)$ and the sender observes $\omega = \emptyset$. In the proposed equilibrium $a^*_i(\emptyset) = (R, \emptyset)$ and this generates $z_1(R, \emptyset) = \xi(1-\widehat{p}_1)$ and $z_2(R, \emptyset) = \xi(\widehat{p}_2)$. Suppose instead the sender deviates to $a_i(\emptyset) = (U, \emptyset)$. This generates a common off-equilibrium belief $\Pr(\theta = \Theta \mid (U, \emptyset)) = \widetilde{p}$, for some $\widetilde{p} \in [0,1]$. This will generate $z_1(U, \emptyset) = \xi(1-\widetilde{p})$ and $z_2(U, \emptyset) = \xi(\widetilde{p})$. By (L), as shown in the proof of Proposition 2.2, $\xi(1-\widetilde{p}) + \xi(\widetilde{p}) \geqslant \xi\left(\frac{1}{2}\right) + \xi\left(\frac{1}{2}\right)$ for all $\widetilde{p} \in [0,1]$. (To see this notice that we can express the $\left(\frac{1}{2}, \frac{1}{2}\right)$ lottery over the original utilities as a lottery with modified probabilities over the expected utilities generated by $(\widetilde{p}, 1-\widetilde{p})$ and $(1-\widetilde{p}, \widetilde{p})$ lotteries on the original utilities and the proof follows the proof of Proposition 2.2). But given our assumption this implies $\xi(1-\widetilde{p}) + \xi(\widetilde{p}) > \xi(1-\widehat{p}_1) + \xi(\widehat{p}_2)$ for all $\widetilde{p} \in [0,1]$. So there exists no off-equilibrium belief that supports the suggested equilibrium.

Suppose $2\xi\left(\frac{1}{2}\right) \leqslant \xi\left(1-\widehat{p}_1\right)+\xi\left(\widehat{p}_2\right)$, then the off equilibrium belief $\Pr\left(\theta=\Theta \mid (U,\emptyset)\right) = \frac{1}{2}$ will prevent the deviation mentioned the previous paragraph. Now suppose the sender observes $\omega=\theta$ for some $\theta \in \{\Theta,-\Theta\}$. In equilibrium $a^*_{\bar{i}(\theta)}(\theta)=(R,\theta)$, $a^*_{\underline{i}(\theta)}(\theta)=(R,\emptyset)$ and thus $z_{\bar{i}(\theta)}(R,\theta)=\xi(1)$ and $z_{\underline{i}(\theta)}(R,\emptyset) \in \{\xi\left(1-\widehat{p}_1\right),\xi\left(\widehat{p}_2\right)\}$. Consider the most profitable deviation, i.e. $a_{\bar{i}(\theta)}(\theta)=r_{\underline{i}(\theta)}(\theta)=(U,\emptyset)$. Since $\xi(1)+\min\{\xi\left(1-\widehat{p}_1\right),\xi\left(\widehat{p}_2\right)\} > \xi\left(1-\widehat{p}_1\right)+\xi\left(\widehat{p}_2\right) \geqslant 2\xi\left(\frac{1}{2}\right)$ the deviation is not profitable. ∎

Lemma 2.5 establishes the necessary and sufficient condition for there to be an equilibrium identical to the one with commitment to private protocol. Notice that in this case, when the sender does not observe any information and announces this privately, this leads to very pessimistic beliefs on the part of both receivers. Hence the sender might be tempted to announce his lack of information in public. In this case the receivers will share the same off-equilibrium belief. Since the receivers are prudent, their choices given these beliefs will constitute a better outcome for the seller than the outcome he would get if the receivers believed both states of the world were equally likely to be true. If this belief performs better for the sender than the pessimistic beliefs, than the deviation beliefs will also perform better. Hence it has to be the case that pessimistic beliefs are preferred by the sender to the belief that the states are equally likely. The next lemma gives sufficient conditions that reverse this preference.

Lemma 2.6. *If* $q > 1-\frac{\min(p,1-p)}{\max(p,1-p)}$, *then* $2\xi\left(\frac{1}{2}\right) > \xi\left(1-\widehat{p}_1\right)+\xi\left(\widehat{p}_2\right)$.

Proof. Suppose $p > \frac{1}{2}$. Then it is easy to see that $\widehat{p}_1 > p > \frac{1}{2}$ and $1-\widehat{p}_1 < \widehat{p}_2 < p < \widehat{p}_1$. The condition in the proposition implies $q < 1-\frac{1-p}{p}$. Given this condition it is easy to see that $\widehat{p}_2 < \frac{1}{2}$. Thus $1-\widehat{p}_1 < \widehat{p}_2 < \frac{1}{2}$. Thus $\xi\left(1-\widehat{p}_1\right) < \xi\left(\widehat{p}_2\right) < \xi\left(\frac{1}{2}\right)$. So it follows that

$\xi(1-\widehat{p}_1)+\xi(\widehat{p}_2)<2\xi\left(\frac{1}{2}\right)$. ∎

Lemma 2.6 says that when the receivers have sufficiently high belief in the fact that the sender is informed, their actions with belief $\left(\frac{1}{2},\frac{1}{2}\right)$ result in a better outcome for the sender than their actions with the pessimistic beliefs. The intuition for this is that, when the probability that the sender receives information is high and he claims that he has not received any information, the receivers will tend to form even more pessimistic beliefs and shade down their action choices further. The following corollary combines this result with Lemma 2.5 and states that when the receivers have high belief in the sender's information, there is no equilibrium where both receivers get partially informed.

Corollary 2.2. *If $q>1-\frac{\min(p,1-p)}{\max(p,1-p)}$, there exists no pure-strategy perfect Bayesian equilibrium with prudent receivers where both receivers get partially informed about opposite signals.*

Having established the non existence of partially informative equilibria for high q, next we establish its existence for sufficiently low q.

Lemma 2.7. *There exists a $0<q^P(p)<1-\frac{\min(p,1-p)}{\max(p,1-p)}$ such that $2\xi\left(\frac{1}{2}\right)\leqslant\xi(1-\widehat{p}_1)+\xi(\widehat{p}_2)$ if and only if $q\leqslant q^P(p)$.*

Proof. Suppose $p>\frac{1}{2}$ (the same argument applies to $p\leqslant\frac{1}{2}$) . Suppose $q=1-\frac{1-p}{p}$. In that case $\widehat{p}_2\left(p,1-\frac{1-p}{p}\right)=\frac{1}{2}$. Let $h(p,q)=\xi(1-\widehat{p}_1(p,q))+\xi(\widehat{p}_2(p,q))$. So $h\left(p,1-\frac{1-p}{p}\right)=\xi\left(1-\widehat{p}_1\left(p,1-\frac{1-p}{p}\right)\right)+\xi\left(\frac{1}{2}\right)<2\xi\left(\frac{1}{2}\right)$ since $1-\widehat{p}_1(p,q)<\frac{1}{2}$ for $p>\frac{1}{2}$. We also know that $\widehat{p}_1(p,0)=\widehat{p}_2(p,0)=p$ and thus $h(p,0)=\xi(1-p)+\xi(p)>2\xi\left(\frac{1}{2}\right)$ by (L) as shown in the proof of proposition 2.2. We also know that $1-\widehat{p}_1(p,q)$ and $\widehat{p}_2(p,q)$ are continuous and strictly decreasing in q. Since $\xi(\widetilde{p})$ is continuous and strictly increasing in $\widetilde{p}$, $h(p,q)$ is continuous and strictly decreasing in q. Thus by intermediate value theorem there exists a $q^P(p)$ such that $h(p,q^P(p)) = \xi(1-\widehat{p}_1(p,q^P(p))) + \xi(\widehat{p}_2(p,q^P(p))) = 2\xi(\frac{1}{2})$ and $h(p,q) \lesseqgtr 2\xi(\frac{1}{2})$ as $q \gtreqless q^P(p)$. ∎

Corollary 2.3. *There exists a pure-strategy perfect Bayesian equilibrium with prudent receivers where both receivers get partially informed about opposite signals if and only if* $q \leqslant q^P(p)$.

We have established the fact the least the prudent receivers could get is partial information and that happens only when $q \leqslant q^P(p)$, i.e. when the probability that the sender gets informed is sufficiently low. The next question is whether there exists a fully-revealing (truthful) equilibrium for both receivers and if so under what conditions. Below we characterize two such equilibria and then show that there exists a truthful equilibrium for all $p, q \in (0,1)$.

Lemma 2.8. *The following are truthful pure-strategy perfect Bayesian equilibria with prudent receivers for all* $p, q \in (0,1)$*:*

(*i*) $a_i^*(\theta) = (R,\theta)$ *and* $a_i^*(\emptyset) = (U,\emptyset)$ *for all* $\theta \in \{\Theta, -\Theta\}$ *and* $i \in I$.

(*ii*) $a_i^*(\theta) = (U,\theta)$, $a_i^*(\emptyset) = (U,\emptyset)$ *for all* $\theta \in \{\Theta, -\Theta\}, i \in I$.

Proof. (*i*) Suppose the sender observes $\omega = \theta \in \{\Theta, -\Theta\}$ as his signal. According to the equilibrium he chooses a private protocol and announces θ to both receivers. So $z_{\bar{i}(\theta)}(a_i^*(\theta)) = \xi(1)$ and $z_{\underline{i}(\theta)}(a_i^*(\theta)) = \xi(0)$. Thus the sender's utility in equilibrium after observing $\omega = \theta$ is $u_s(\xi(1) + \xi(0))$. Now, let us consider the possible deviations for the sender. One possible deviation is $a_{\underline{i}(\theta)}(\theta) = (R,\emptyset)$ or $a_{\bar{i}(\theta)}(\theta) = (R,\emptyset)$. Assign the following off-equilibrium belief to receiver $\underline{i}(\theta)$: $\Pr(\theta \mid (R, \emptyset)) = 1$. Given this belief, $z_{\underline{i}(\theta)}(R, \emptyset) = \xi(0)$. So the deviation of this sort is not profitable. Another possible deviation is $a_{\underline{i}(\theta)}(\theta) = a_{\bar{i}(\theta)}(\theta) = (U, \theta)$. But this is equivalent to $a^*_{\underline{i}(\theta)}(\theta) = a^*_{\bar{i}(\theta)}(\theta) = (R, \theta)$. Next possible deviation is $a_{\underline{i}(\theta)}(\theta) = a_{\bar{i}(\theta)}(\theta) = (U, \emptyset)$. In this case $z_1(U, \emptyset) = \xi(1-p)$ and $z_2(U, \emptyset) = \xi(p)$, thus the sender's utility is $u_s(\xi(1-p) + \xi(p))$. By equation (2.23), $u_s(\xi(1-p) + \xi(p)) < u_s(\xi(0) + \xi(1))$, so this deviation is not profitable either. The last deviation we should consider is $a_i(\emptyset) = (R, \emptyset)$ for all $i \in I$. Because of the off-equilibrium belief we assumed above $z_i(R, \emptyset) = \xi(0)$ for all $i \in I$. Again in this case the sender's utility is $u_s(2\xi(0)) < u_s(\xi(1-p) + \xi(p))$. So no deviations are profitable given the off-equilibrium behavior we assumed on the part of the receivers.

(ii) For any $\theta \in \{\Theta, -\Theta\}$ let us assume the following off-equilibrium belief for receiver $\underline{i}(\theta) : \Pr(\theta \mid (R, \emptyset)) = 1$. Given this off equilibrium belief, upon observing $\omega = \theta$ $a^*_i(\theta) = (U, \theta)$ and the possible deviation $a_{\bar{i}(\theta)}(\theta) = \left(R, \widehat{\theta}\right)$, $a_{\underline{i}(\theta)}(\theta) = (R, \emptyset)$ generate the same outcome, namely $z_{\bar{i}(\theta)}(U, \theta) = z_{\bar{i}(\theta)}(R, \theta) = \xi(1)$ and $z_{\underline{i}(\theta)}(U, \theta) = z_{\underline{i}(\theta)}(R, \emptyset) = \xi(0)$. So such a deviation is not strictly profitable. We already now that no profitable deviations within the public protocol exist because we showed before that truthfulness was an equilibrium. It is trivial to see that no other profitable deviation exists. ∎

Both equilibria are truthful. The first one employs a combination of private and public protocols for full revelation, whereas the second one is equivalent to the truthful equilibrium we found when the seller could commit to the public protocol. Especially the first equilibrium illustrates the importance of awareness by the receivers that the sender has a public announcement opportunity available to him. Our truthful equilibria imply that when the sender goes to the prudent receiver and reports that he has no information, the receiver does not believe it as it is but asks the question "why does he not announce this on TV?."

Proposition 2.5. *There exists a truthful equilibrium for all* $q, p \in (0,1)$ *and all the equilibria are truthful when* $q \geqslant q^P(p)$.

Proof. Follows from all the previous lemmas and corollaries. ∎

We see that when the receivers are prudent there is multiple equilibria, but truthfulness is always an equilibrium whereas partial revealing is only equilibrium for a range of parameter values. Notice that since $q^P(p) < 1 - \frac{\min(p,1-p)}{\max(p,1-p)}$, as p gets closer to $\frac{1}{2}$, the likelihood of the partial revealing equilibrium decreases. If we think that p and q are chosen randomly, we would expect to see the truthful equilibrium more often than the partial revealing one. It is true that commitment to public protocol drives full truthfulness, but it seems like the revealing power of public protocol is not whitewashed even when we do not allow for commitment.

2.4.2 Equilibria with haphazard receivers

In this section we investigate information transmission to haphazard receivers when the sender cannot commit to a protocol. We follow the same steps as in the case with prudent receivers. Since haphazard receivers are defined as the exact opposites of prudent receivers, the results here almost constitute a mirror image of the results in the previous section. In particular we first show that in equilibrium the sender always prefers to withhold some information to revealing it all. Then we show that in equilibrium, regardless of the signal, it can not be the case that one receiver gets more informed than the other. Using these results we show that there are two possible equilibria, one where each receiver gets informed about only one signal and another one where they are left uninformed about both signals. Similar to the previous section, we find that the first kind of equilibrium exists only for a restricted set of parameters, whereas the second one always exists.

Lemma 2.9. *There exists no pure-strategy perfect Bayesian equilibrium that is truthful with haphazard receivers, i.e. there exists no equilibrium where for all* $\theta \in \{\Theta, -\Theta\}$, *for all* $i \in I$ *and for some* $\gamma \in \Gamma$ $a_i^*(\theta) = (\gamma, \theta)$.

Proof. For all θ in equilibrium the sender gets $u_s(\xi(0) + \xi(1))$. Consider the deviation $a_i(\theta) = a_i^*(\emptyset)$. Given the truthful equilibrium, $z_1(a_i^*(\emptyset)) = \xi(1-p)$, $z_2(a_i^*(\emptyset)) = \xi(p)$. Thus the sender's utility with the deviation is $u_s(\xi(1-p) + \xi(p)) > u_s(\xi(0) + \xi(1))$ by equation (2.24) ∎

Lemma 2.9 establishes the fact that when the receivers are haphazard, the sender always prefers to hide some information. Hence there exists no truthful equilibrium.

Lemma 2.10. *There exists no pure-strategy perfect Bayesian equilibrium with haphazard receivers where one receiver in equilibrium gets more information than the other, i.e. in all equilibria if there exists* $\theta, \theta' \in \{\Theta, -\Theta\}$ *and* $\omega \in \{\theta', \emptyset\}$ *such that* $a_{\bar{i}(\theta)}^*(\theta) = (R, \theta)$, $a_{\underline{i}(\theta)}^*(\theta) = a_{\underline{i}(\theta)}^*(\omega) = (R, \emptyset)$ *then* $a_{\underline{i}(\theta)}^*(\theta') = (\gamma, \theta')$ *for some* $\gamma \in \Gamma$ *and* $a_{\bar{i}(\theta)}^*(\theta') = a_{\bar{i}(\theta)}^*(\emptyset)$.[2]

Proof. The same as the proof of the same lemma in the previous section. ∎

[2]Notice that the non-existence of such an equilibrium with $\underline{i}(\theta)$ as the more informed follows from the arguement in the prudent receivers case.

The previous two lemmas show that the two receivers get at most partially informed and they should get informed in exactly the same number of cases. First let us look at the case where they both get partially informed. Notice that the case where the receivers get partially informed about the same signal is trivial. Suppose they both get informed in private or in public when $\omega = \theta$, but not when $\omega = \theta'$ where $\theta,\ \theta' \in \{\Theta, -\Theta\}$, i.e. $a_i^*(\emptyset) = a_i^*(\theta')$ and $a_i^*(\emptyset) \neq a_i^*(\theta')$ for all $i \in I$. Then when $\omega = \theta$, the receivers responses will be $\xi(0) + \xi(1)$. But then the sender would rather deviate to $a_i(\theta) = a_i^*(\emptyset)$, because by equation (2.24), $\min\{\xi(1-\widehat{p}_1) + \xi(\widehat{p}_1), \xi(1-\widehat{p}_2) + \xi(\widehat{p}_2)\} > \xi(0) + \xi(1)$. Having talked about this case, the next case is where they get partially informed about opposite signals. We will next look for such an equilibrium, which in fact is identical to the equilibrium we had in the case where the sender commits to private protocol.

Lemma 2.11. *There exists a pure-strategy perfect Bayesian equilibrium with haphazard receivers where both receivers get partially informed about opposite signals, i.e. an equilibrium where* $a_{\bar{i}(\theta)}^*(\theta) = (R, \theta)$, $a_{\underline{i}(\theta)}^*(i) = (R, \emptyset)$ *and* $a_i^*(\emptyset) = (R, \emptyset)$ *for all* $i \in I$ *, if and only if* $\xi(0) + \xi(1) \leqslant \xi(1-\widehat{p}_1) + \xi(\widehat{p}_2)$

Proof. Suppose $\xi(0) + \xi(1) > \xi(1-\widehat{p}_1) + \xi(\widehat{p}_2)$ and the sender observes $\omega = \emptyset$. In the proposed equilibrium $a_i^*(\emptyset) = (R, \emptyset)$ and this generates $z_1(R, \emptyset) = \xi(1-\widehat{p}_1)$ and $\xi_2(R, \emptyset) = \xi(\widehat{p}_2)$. Suppose instead the sender deviates to $a_i^*(\emptyset) = (U, \emptyset)$. This generates a common off-equilibrium belief $\Pr(\theta = \Theta \mid (U, \emptyset)) = \widetilde{p}$, for some $\widetilde{p} \in [0,1]$. This will generate $z_1(U, \emptyset) = \xi(1-\widetilde{p})$ and $z_2(U, \emptyset) = \xi(\widetilde{p})$. By (G), $\xi(1-\widetilde{p}) + \xi(\widetilde{p}) > \xi(0) + \xi(1)$ for all $\widetilde{p} \in [0,1]$. But given our assumption this implies $\xi(1-\widetilde{p}) + \xi(\widetilde{p}) > \xi(1-\widehat{p}_1) + \xi(\widehat{p}_2)$ for all $\widetilde{p} \in [0,1]$. So there exists no off-equilibrium belief that supports the suggested equilibrium.

Suppose $\xi(0)+\xi(1) \leqslant \xi(1-\widehat{p}_1)+\xi(\widehat{p}_2)$. The off equilibrium belief $\Pr(\theta = \Theta \mid (U, \emptyset)) = 1$ will prevent the deviation mentioned the previous paragraph. Now suppose the sender observes $\omega = \theta$ for some $\theta \in \{\Theta, -\Theta\}$. In equilibrium $a^*_{\bar{i}(\theta)}(\theta) = (R, \theta)$, $a^*_{\underline{i}(\theta)}(\emptyset) = (R, \emptyset)$ and thus $z_{\bar{i}(\theta)}(R, \theta) = \xi(1)$ and $z_{\underline{i}(\theta)}(R, \emptyset) \in \{\xi(1-\widehat{p}_1), \xi(\widehat{p}_2)\}$. Since $\xi(1) + \min\{\xi(1-\widehat{p}_1), \xi(\widehat{p}_2)\} > \xi(1-\widehat{p}_1) + \xi(\widehat{p}_2) \geqslant \xi(0) + \xi(1)$ and thus the deviation is not profitable. ∎

Lemma 2.11 gives the necessary and sufficient conditions for the existence of a partially revealing equilibrium, which is equivalent to the equilibrium when the sender commits to a private protocol. The intuition is as follows: If the sender does not observe any information and announces this privately, the receivers form pessimistic beliefs about the state of the world. The sender can alternatively deviate and announce his lack of information publicly. This generates some off-equilibrium belief. However, since the receivers are haphazard, their total response is at least as high with the off-equilibrium beliefs as with perfect information. If the sender prefers the perfect information response to the pessimistic belief response, then he also prefers off-equilibrium response to the pessimistic belief response and hence he deviates. Therefore, it has to be the case that the sender prefers pessimistic beliefs to perfect information. The next lemma provides necessary and sufficient conditions for this preference to hold.

Lemma 2.12. *There exists a* $0 < q^H(p) < 1$ *such that* $\xi(0) + \xi(1) \leqslant \xi(1-\widehat{p}_1) + \xi(\widehat{p}_2)$ *if and only if* $q \leqslant q^H(p)$.

Proof. Let $h(p,q) = \xi(1-\widehat{p}_1(p,q)) + \xi(\widehat{p}_2(p,q))$. We know that $h(p,0) = \xi(1-\widehat{p}_1(p,0)) + \xi(\widehat{p}_2(p,0)) = \xi(1-p) + \xi(p) > \xi(0) + \xi(1)$ by (G). We also know that $h(p,1) =$

$\xi(1-\widehat{p}_1(p,1))+\xi(\widehat{p}_2(p,1))=2\xi(0)<\xi(0)+\xi(1)$. Since $\xi(.)$ is continuous and strictly increasing , $h(p,q)$ is continuous and strictly decreasing in q. Thus by intermediate value theorem there exists a $q^H(p)$ such that $h(p,q^H(p))=\xi(1-\widehat{p}_1(p,q^H(p)))+\xi(\widehat{p}_2(p,q^H(p)))=\xi(0)+\xi(1)$ and $h(p,q)\lesseqgtr\xi(0)+\xi(1)$ as $q\gtreqless q^H(p)$. ∎

Lemma 2.12 says that the sender prefers the pessimistic beliefs only when there is low probability that he gets informed. The reason is that when that probability is sufficiently low, the receivers hold less pessimistic beliefs and hence choose higher actions. Using this result and Lemma 2.11 the following corollary concludes that q needs to be low for partially revealing equilibrium to exist.

Corollary 2.4. *There exists a pure-strategy perfect Bayesian equilibrium with haphazard receivers where both receivers get partially informed about opposite signals if and only if* $q\leqslant q^H(p)$.

We have established the fact that the most the haphazard receivers could get is partial information and that happens only when $q\leqslant q^H(p)$, i.e. when the probability that the sender gets informed is sufficiently low. The next question is whether there exists a non-revealing equilibrium for both receivers and if so under what conditions.

Proposition 2.6. *There exists a unique non-revealing equilibrium for all* $p,q\in(0,1)$. *It is characterized by* $a_i^*(\omega)=(U,\emptyset)$ *for all* $\omega\in\Omega$ *and all* $i\in I$. *It is the unique equilibrium when* $q\geqslant q^H(p)$.

Proof. Let us first show that $a_i^*(\omega)=(U,\emptyset)$ is an equilibrium for all $p,q\in(0,1)$. Because of (G) we know that $\xi(0)+\xi(1)<\xi(1-p)+\xi(p)$, thus the deviations of the sort $a_i(\theta)=(U,\theta)$

are ruled out. For all $\theta \in \{\Theta, -\Theta\}$ consider the off-equilibrium belief for player $\bar{\imath}(\theta)$ to be $\Pr(\theta \mid (R, \emptyset)) = 0$. Consider the most profitable deviation $a_{\bar{\imath}(\theta)}(\theta) = (R, \theta)$ and $a_{\hat{\imath}(\theta)}(\theta) = (R, \emptyset)$. With this deviation and the off-equilibrium belief the sender would get at most $\xi(1) + \max\{\xi(1-\widehat{p}), \xi(\widehat{p}_2)\} = \xi(1) + \xi(0) < \xi(1-p) + \xi(p)$ by equation (2.24). So there exists no profitable deviations and $a_i^*(\omega) = (U, \emptyset)$ is an equilibrium.

Suppose there is another non-revealing equilibrium. It has to be the case that for some $\gamma \in \Gamma$ and for all $i \in I$, $a_i^*(\omega) = (\gamma, \emptyset)$. If $\gamma \neq U$ then $\gamma = R$. So, in this equilibrium the sender gets $u_s(\xi(1-p) + \xi(p))$. But if for some $\theta \in \{\Theta, -\Theta\}$ the sender deviates to $a_{\bar{\imath}(\theta)}(\theta) = (R, \theta)$, $z_{\bar{\imath}(\theta)}(R, \theta) = \xi(1)$ and so he would deviate. Thus there exists no other non-revealing equilibrium.

The fact that it is the unique equilibrium when $q \leqslant q^H(p)$ follows from Lemmas 2.11 and 2.12 and Corollary 2.5. ∎

There is multiple equilibria with haphazard receivers too, but non-revealing is always an equilibrium whereas partial revealing is only equilibrium for a range of parameter values. Again if we think that p and q are chosen randomly, we would expect to see the non-revealing equilibrium more often than the partial revealing one. Similar to what we saw with prudent receivers, here again the power of public protocol, though not as strong as it was when there was commitment, still exists.

Looking at the interaction with both the haphazard and the prudent receivers we see that the possibility of a public announcement tends to dominate the interaction. In other words, the equilibrium we find without commitment is closer to the equilibrium we find when the sender commits to public protocol. We believe this is very interesting, because although we physically have the chance to whisper to people's ears, the fact that we can always shout out loud on the street makes the whispering ineffective most of the time. Returning to the politician example, our results imply that as long as the media makes it possible for the politician to make a public announcement at any instant, even behind closed doors, the politician can not diverge much from what he would say on a TV announcement.

2.5 Extension to Unlimited Protocols with No Commitment

As we mentioned in the introduction, the senders are usually not restricted to make one-shot announcements. Here we follow the no commitment framework in the previous section, but we extend the model to allow the sender to make as many announcements as he likes in the protocols of his choice until some deadline. For example, the politicians usually make a lot of private meetings with different groups followed by a public announcement, or vice versa. The deadline for the politician can be considered as the election day.

We will model this in the following way. Suppose the receivers start listening to the sender at $t = 0$ and make their choices at $t = 1$. The sender no longer is limited to make one announcement to each group. He still has to make at least one announcement to each receiver in private or public, but after the first announcement he can make as many more as he wants till just before $t = 1$ in the protocols of his choice. Here when the sender makes an announcement to receiver i in private, not only does receiver i' not observe what the announcement is, but also he is not aware of such a private announcement being made to receiver i. (In the previous case where we had the "one and only one announcement to each receiver" rule, making this distinction would not change anything because one announcement to each receiver was required and thus when receiver i' was told something in private, he could infer that receiver i also received a private announcement).

Here the sender's decision at each $t \in [0,1)$ is whether to make an announcement to each receiver i and if so in what protocol and what announcement to make contingent on all the past protocol choices and announcements and his signal $\omega \in \Omega$. Again at each stage we are going the assume the protocol decision and the announcement decision as a joint one. So at time t the sender chooses a protocol $\gamma = (\gamma_1, \gamma_2) \in \Gamma = [\{R_1, N_1\} \times \{R_2, N_2\}] \cup \{U, U\}$ where R_i stands for a private meeting with receiver i and N_i stands for no announcement to receiver i. Together with the protocol he chooses an announcement to receiver i if the protocol he chose necessitates an announcement to receiver i. The announcement to receiver i is given by $\omega_i \in \Omega$. The history h_t at time t consists of past choices $(\gamma_\tau, \omega_{1\tau}, \omega_{2\tau})_{\tau<t}$ and $h_0 = \{\}$. Let h_{it} be the history up to time t that is observable to receiver i. Let $\Gamma_i = \{\gamma_i \mid (\gamma_i, \gamma_{i'}) \in \Gamma \text{ for some } \gamma_{i'}\}$ and H_t be the set of all possible histories. The sender's strategy then is $a_{it} : \Omega \times H_t \to \Gamma_i \times \Omega$ for $i \in I$ and $t \in [0,1)$

Unlimited Protocol Verifiability Assumption: For all $\omega \in \Omega$, $i \in I$, $t \in [0,1)$ and $h_t \in H_t$, $a_{it}(\omega, h_t) \in \{(\gamma_i, \widehat{\omega}) \mid \gamma_i \in \Gamma_i \text{ and } \widehat{\omega} \in \{\omega, \emptyset\}\}$.

Unlimited Protocol Uniform Announcement Assumption : For all $\omega \in \Omega$ and $h_t \in H_t$ for which $\exists i \in I$ such that $a_i(\omega, h_t) \in \{(U, \widehat{\omega}) \mid \widehat{\omega} \in \Omega\}$, $a_{it}(\omega, h_t) = a_{i't}(\omega, h_t)$.

At least one announcement requirement: For all $\omega \in \Omega$, $i \in I$ there exists $t \in [0,1)$ such that $a_{it}(\omega, h_t) \in \{(\gamma_i, \widehat{\omega}) \mid \gamma_i \in \Gamma_i / \{N_i\} \text{ and } \widehat{\omega} \in \Omega\}$.

The first assumption is the verifiable reports assumption, the second one means the two receivers receive the same announcement in a public protocol and the third one means each receiver receives at least one announcement.

Lemma 2.13. *Let a_{it}^* represent the equilibrium protocol and announcement choices and $h_t^*(\omega)$ the history generated under equilibrium play as a function of sender's information $\omega \in \Omega$. There is no pure strategy perfect Bayesian equilibrium where for some $\theta \in \{\Theta, -\Theta\}$ and some $\omega \in \Omega/\{\theta\}$ $a_{\bar{\imath}(\theta)t}^*(\theta, h_t^*(\theta)) = a_{\bar{\imath}(\theta)t}^*(\omega, h_t^*(\phi))$ for all $t \in [0,1)$. In other words in all equilibria when the seller gets to observe θ, receiver $\bar{\imath}(\theta)$ gets to infer it somewhere along the way.*

Proof. Suppose there was such an equilibrium. Then there exists a $\theta \in \{\Theta, -\Theta\}$ such that at $t = 1$ receiver $\bar{\imath}(\theta)$ has a belief $\Pr(\theta = \theta \mid h_{it}^*) < 1$, because there exists a $\omega \in \Omega/\{\theta\}$ such that $h_{\bar{\imath}(\theta)t}^*(\theta) = h_{\bar{\imath}(\theta)t}^*(\omega)$ for all $t \in [0,1)$. So $z_{\bar{\imath}(\theta)}\left(h_{\bar{\imath}(\theta)1}^*\right) < \xi(1)$. Receiver $\underline{i}(\theta)$'s equilibrium choice is $z_{\underline{i}(\theta)}\left(h_{\underline{i}(\theta)1}^*\right)$. The sender gets $u_s\left(z_{\bar{\imath}(\theta)}\left(h_{\bar{\imath}(\theta)1}^*\right) + z_{\underline{i}(\theta)}\left(h_{\underline{i}(\theta)1}^*\right)\right) < u_s\left(\xi(1) + z_{\underline{i}(\theta)}\left(h_{\underline{i}(\theta)1}^*\right)\right)$ in equilibrium. But then the sender would actually prefer to deviate to $\lim_{\tau \to 1} a_{\bar{\imath}(\theta)\tau}\left(\widehat{\theta}, h_\tau^*(\theta)\right) = \left(R_{\bar{\imath}(\theta)}, \theta\right)$ thus inducing receiver $\bar{\imath}(\theta)$ to choose $\xi(1)$ and guaranteeing $u_s\left(\xi(1) + z_{\underline{i}(\theta)}\left(h_{\underline{i}(\theta)1}^*\right)\right)$. ∎

Lemma 2.13 says that the sender always manages to convey the right information to the right receiver. The logic behind the proof is that, if the sender observes θ, he can always add a private meeting with receiver $\bar{\imath}(\theta)$ just before $t = 1$ and inform him about the true information without changing the response of receiver $\underline{i}(\theta)$.

Lemma 2.14. *Let a_{it}^* represent the equilibrium protocol and announcement choices and $h_t^*(\omega)$ the history generated under equilibrium play as a function of sender's information $\omega \in \Omega$. There is no pure strategy Perfect Bayesian equilibrium where there exists an $\theta \in$*

$\{\Theta, -\Theta\}$ *and* $t \in [0,1)$ *such that* $a^*_{\underline{i}(\theta)t}(\theta, h^*_t(\theta)) \neq a^*_{\underline{i}(\theta)t}(\omega, h^*_t(\omega))$ *for all* $\omega \in \Omega/\{\theta\}$. *In other words in all equilibria when the seller gets to observe* θ, *he hides it from receiver* $\underline{i}(\theta)$.

Proof. Suppose there is such an equilibrium where receiver $\underline{i}(\theta)$ gets to hear it when the seller observes $\omega = \theta \in \{\Theta, -\Theta\}$. So receiver $\underline{i}(\theta)$ in equilibrium chooses $z_{\underline{i}(\theta)}\left(h^*_{\underline{i}(\theta)1}(\theta)\right) = \xi(0)$. By Lemma 2.13 we know in this equilibrium for some $t \in [0,1)$ and $a^*_{\bar{i}(\theta)t}(\theta, h^*_t(\theta)) \neq a^*_{\bar{i}(\omega)t}(\omega, h^*_t(\omega))$ for all $\omega \in \Omega/\{\theta\}$. Thus receiver $\bar{i}(\theta)$'s choice in equilibrium will be $z_{\bar{i}(\theta)}\left(h^*_{\bar{i}(\theta)1}\right) = \xi(1)$. So the sender in equilibrium gets $u_s(\xi(1) + \xi(0))$. Suppose the sender deviates to $a^*_{\bar{i}(\theta)t}(\emptyset, h^*_t(\emptyset))$, $a^*_{\underline{i}(\theta)t}(\emptyset, h^*_t(\emptyset))$ for all t and $\lim_{\tau \to 1} a_{\bar{i}(\theta)\tau}(\theta, h^*_\tau(\emptyset)) = \left(R_{\bar{i}(\theta)}, \theta\right)$. With this deviation $z_{\underline{i}(\theta)}\left(h^*_{\underline{i}(\theta)1}(\emptyset)\right) > \xi(0)$ because for $\underline{i}(\theta)$ it induces the belief $\Pr\left(\omega = \emptyset \mid h^*_{\underline{i}(\theta)1}(\emptyset)\right) > 0$ and $z_{\bar{i}(\theta)}\left(\left[h^*_{\bar{i}(\theta)1}(\emptyset), \left(R_{\bar{i}(\theta)}, \theta\right)\right]\right) = \xi(1)$. Thus the seller gets $u_s\left(\xi(1) + z_{\underline{i}(\theta)}\left(h^*_{\underline{i}(\theta)1}(\emptyset)\right)\right) > u_s(\xi(1) + \xi(0))$. ∎

Lemma 2.14 says that the sender manages to hide the information from the receiver who does not want to hear it. The logic behind this proof is that the sender can always imitate the equilibrium behavior of a sender who has no information and then have a private meeting with the receiver who wants to hear the information.

Lemmas 2.13 and 2.14 make it clear that it is not possible to have a truthful or non-revealing equilibrium. In other words, all equilibria are partially revealing, i.e. equivalent to the equilibrium in the case of commitment to the private protocol. However, in the previous two sections we showed that both with prudent and haphazard receivers such an equilibrium exists only for low levels of q. The next proposition shows that this result holds for the unlimited protocol case as well.

Proposition 2.7. *There exists a pure strategy perfect Bayesian equilibrium with prudent* *receivers if and only if $q \leqslant q^P(p)$ and there exists a pure strategy Bayesian equilibrium with haphazard receivers if and only if $q \leqslant q^H(p)$ where $q^P(p)$ and $q^H(p)$ are as defined before. If $q \leqslant q^k(p)$ for $k \in \{P, H\}$ then the following is an equilibrium with k type receivers and all other equilibria result in the same inference and payoffs: For all $\theta \in \{\Theta, -\Theta\}$ and $i \in I$, $a^*_{\bar{i}(\theta)0}(\theta, \{\}) = \left(R_{\bar{i}(\theta)}, \theta\right)$, $a^*_{\underline{i}(\theta)0}(\theta, \{\}) = \left(R_{\underline{i}(\theta)}, \emptyset\right)$, $a^*_{i0}(\emptyset, \{\}) = (R_i, \emptyset)$ and $a^*_{it}(\omega, h^*_t(\omega)) = (N_i, .)$ for all $t \in (0,1)$ and $\omega \in \Omega$. In other words the equilibrium is equivalent to the one we had in the private protocol case.*

Proof. Take prudent receivers. Let a^*_{it} represent the equilibrium protocol and announcement choices and $h^*_t(\omega)$ the history generated under equilibrium play as a function of sender's information $\omega \in \Omega$. By the previous two lemmas we know that for all $\theta \in \{\Theta, -\Theta\}$ there exists a $\widehat{t} \in (0,1)$ such that $a^*_{\bar{i}(\theta)\widehat{t}}\left(\theta, h^*_{\widehat{t}}(\theta)\right) = \left(R_{\bar{i}(\theta)}, \theta\right)$ and there exists a $\omega \in \Omega / \{\theta\}$ such that for all $t \in (0,1)$ $a^*_{\underline{i}(\theta)t}(\theta, h^*_t(\theta)) = a^*_{\underline{i}(\theta)t}(\omega, h^*_t(\omega))$. In particular since due to the last two lemmas $\underline{i}(\theta)$ gets informed of $\theta' = \{\Theta, -\Theta\} / \{\theta\}$, it has to be the case that $\omega = \emptyset$. Thus $a^*_{\underline{i}(\theta)\widehat{t}}\left(\theta, h^*_{\widehat{t}}(\theta)\right) = a^*_{\underline{i}(\theta)\widehat{t}}\left(\emptyset, h^*_{\widehat{t}}(\emptyset)\right) \neq (U, \emptyset)$, and it follows that $a^*_{\bar{i}(\theta)\widehat{t}}\left(\emptyset, h^*_{\widehat{t}}(\emptyset)\right) \neq (U, \emptyset)$. So $z_{\bar{i}(\theta)}\left(h^*_{\bar{i}(\theta)1}(\theta)\right) = \xi(1)$ and $z_{\underline{i}(\theta)}\left(h^*_{\underline{i}(\theta)1}(\theta)\right) \in \{\xi(1-\widehat{p}), \xi(\widehat{p}_2)\}$. Alternatively the sender could deviate to $a_{i\widehat{t}}\left(\theta, h^*_{\widehat{t}}(\theta)\right) = (U, \emptyset)$ where $\widehat{t}$ is as formerly defined. Suppose this generates an off-equilibrium belief $(\widetilde{p}, 1-\widetilde{p})$. So the sender would get $u_s(\xi(\widetilde{p}) + \xi(1-\widetilde{p}))$. But then this is a problem that we have already analyzed in Corollary 2.3 and 2.4 for both receivers in the previous section. Such an off-equilibrium belief that would not cause a deviation exists if and only if $q \leqslant q^P(p)$ for the prudent receivers and $q \leqslant q^H(p)$ for the haphazard receivers. Thus there exists no pure strategy perfect Bayesian equilibrium if these conditions are violated.

Now suppose $q \leqslant q^P(p)$ and $q \leqslant q^H(p)$, we need to show that $a^*_{\bar{i}(\theta)0}(\theta, \{\}) = \left(R_{\bar{i}(\theta)}, \theta\right)$, $a^*_{\underline{i}(\theta)0}(\underline{i}(\theta), \{\}) = \left(R_{\underline{i}(\theta)}, \emptyset\right)$, $a^*_{i0}(\emptyset, \{\}) = (R_i, \emptyset)$, $a^*_{it}(\omega, h^*_t(\omega)) = (N_i, .)$ constitute an equilibrium. We have already checked for the deviation $a_{it}(\theta, h^*_t(\theta)) = (U, \emptyset)$ for any $t \in [0, 1)$. For the other possible deviations, i.e. the deviations $a_{it}(\theta, h^*_t(\theta)) = (\gamma_i, .)$ where $\gamma_i \in \Gamma / \{U\}$, for all $t \in [0, 1)$ let the off-equilibrium belief of receiver $\bar{i}(\theta)$ be $\Pr(\theta \mid h^*_t(\theta), a_{it}(\theta, h^*_t(\theta))) = 0$. It is clear that with such an off-equilibrium belief there cannot be any profitable deviations.

The equivalence of payoffs and inference among all equilibria is due to the previous two lemmas. ∎

Adding the possibility of further announcements changed the picture completely. Contrary to the one-shot no commitment case, here we lost the pure strategy equilibria arising from the power of existence of a public announcement option, i.e. the truthful equilibria with prudent receivers and the non-revealing equilibria with haphazard receivers. The fact that the sender can always have one last secret meeting with a receiver dominates the interaction. Our result is not very powerful in the sense that it can not characterize all the equilibria (possible mixed strategy equilibria). However, it still shows that when it is not very likely that the sender is more informed than the receivers, the possibility of making hundreds of announcements is equivalent to committing to a private protocol in the beginning.

2.6 Conclusion

We showed that the information transmission problem is very different with multiple audiences than with a single audience. We introduced a two receiver model that restricts the two receivers to have exactly opposite interests and the reports to be verifiable. We first investigated the informativeness of equilibria when the sender can commit to either private or protocol case. We found that the degree of revelation depends on the receivers' preferences. In particular we categorized the receivers as prudent and haphazard according to the differences in their responses to good and bad news. We showed that the private protocol is more informative with haphazard receivers and public protocol is more informative with prudent receivers.

We also analyzed the equilibria when the sender cannot commit to a particular protocol. We showed that commitment to public protocol is sufficient but not necessary for the interest conflict between the two receivers to affect the communication. In other words, even when the sender cannot commit, common knowledge that the sender has the option of a public announcement is powerful enough to create the same affect as a committed public announcement. However, following this we showed that this power disappears when the sender is not restricted to make a one-shot announcement. The receivers can never be sure whether what they hear now will be the last thing they hear. In other words, as the sender spends more time and effort on making many announcements about the same issue, he can always find one particular instant where only one receiver is present and thus he can always give out the information to the receiver who will like it. This makes the private interactions more powerful in determining the informativeness of the communication. This provides a possible explanation to why, in the politician's case, there are still closed door meetings, why each group of constituent knows something that the others do not know and in particular why the politician gets away making promises to different groups without the others knowing.

Our model is restrictive because it assumes that the receivers have exactly opposite interests. This assumption allows us to highlight the differences between private and public protocols. It would be interesting to investigate a model that allows for different levels of correlation between the receivers interest. As the interests of the receivers converge, we expect the equilibria in private and public protocols to converge as well. The unlimited protocols case constitutes a topic for further research as well. It would be interesting to see whether the results here could be generalized to an informationally richer model that allows for cheap talk.

Chapter 3

Dynamic Monopoly With Variable Quality

3.1 Introduction

Dynamic monopoly problems have been analyzed widely in the literature especially after the well known "Coase conjecture". Coase argued in 1972 that if a durable good monopolist sells over time and if he can make offers very rapidly, he will open the market almost at marginal cost. The argument relies on the fact that after each sale there will be a residual demand and the monopolist will lower the price each time towards marginal cost and move down the demand curve. This argument is surprising because it implies that a dynamic monopolist generates the competitive outcome.

The formal treatments of this conjecture include Bulow (1982), Stokey (1982), Sobel and Takashi (1983) and Fudenberg *et al.* (1985) among others. The most elaborate analysis of Coase conjecture was by Gul, Sonnechein and Wilson (1986). They investigate two different cases: (1) "the gap case" where the lowest valuation in the market is above marginal cost and (2) "the no gap case" where the lowest valuation is below marginal cost. The gap case implies that the market will be covered in finite time. When the offers are made very quickly, the market reaches the final price very fast. This means that the buyers can get lower prices very soon which makes them unwilling to accept much higher offers. Therefore the market has to open with a price sufficiently close to the lowest price. The no gap case is trickier because the market can stay open forever and uniqueness is not guaranteed. In that case Gul, Sonnechein and Wilson (1986) verify the Coase conjecture for a subset of stationary equilibrium.

The proofs given for Coase conjecture in the papers mentioned above explicitly assume the following: (1)durability, (2)no resale or rent option, (3)no capacity constraint and (4)stationarity of equilibrium. The conjecture's robustness to these assumptions has been questioned in the literature. For example Ausubel and Deneckere (1989) show that almost anything can be supported in the no gap case in a non-stationary equilibrium by using the stationary equilibrium in Gul et al. (1986) as a threat. Recently McAfee and Wiseman (2004) show that if there is a cost of increasing the capacity, it acts as a commitment device for the monopolist and the monopolist can enjoy profits bounded away from zero.

There is also an implicit assumption behind the conjecture and all its verifications. The assumption is that the good that is sold is the same every period. Karp (1996) introduces an exogenous variation over time, namely depreciation of the good. He shows that there are some Strong Markov Perfect Equilibria which entail steady-state production at a level lower than in the competitive equilibrium and hence the seller makes positive profits. There are other scenarios for exogenous change, but in this paper, we take the analysis a step further and consider endogenous variation of the good. There are many instances in real life where the seller decides to change the characteristics of the product over time. As a most recent example, consider the very popular audio product by Apple, namely iPod. There are many different versions of iPod, but let us focus on the lowest end product, the so-called mini iPod. After almost a year Apple had introduced mini iPod, they introduced another version with higher memory. So the buyers who still have not purchased an iPod now face one with higher quality. In this case we could think that Apple is walking its way up the ladder in terms of buyers' willingness to pay. There are examples for the opposite kind of move. Consider Toshiba and HP in the laptop market. These companies are known to introduce high quality products first and then lower the quality once they realize they are not meeting their sales target. In this category, Dell constitutes an example that walks up the ladder like Apple's iPods.

All these examples highlight the fact that even in the most abstract world of a dynamic monopolist, it is natural to expect the monopolist to vary the characteristics of the product over time. In this paper, we consider variation in the quality of the product. We take the standard dynamic monopolist problem and we add the possibility of quality variation each period, i.e. the seller announces a new quality level together with a new price each period. The buyers are differentiated with respect to how much they value quality. Notice that this can also be interpreted as the seller offering a menu of contracts over time. In other words, this is equivalent to the problem of a seller who is restricted to make one offer at once, but can use time as a way to separate buyers with different willingness to pay and hence extract more surplus. This "adverse selection" aspect of the problem introduces an interesting dynamic into the model. If the buyers think that they will get a low price in the future, they must also believe that they will get a low quality. This can be expected to dampen the buyers' willingness to wait. We investigate whether this dampening is enough to overcome Coase conjecture.

As it is evident from the examples above, it is not clear whether the seller will move up or down the demand curve over time. We show that with finite number of periods the latter holds in equilibrium, i.e. the seller serves the buyers with higher values earlier. Then we look at the infinite period equilibrium which constitutes the limit for the finite period game. This equilibrium is stationary and it inherits the downward movement over the demand curve. We show that in this equilibrium Coase conjecture is verified, however not in its usual sense. The seller's profit goes to zero as the offers are made more frequently, but not through zero prices. In the limit the seller offers the efficient menu of contracts and leaves all the surplus to the buyers.

The road map is as follows: We describe the model in Section 3.2. We first analyze the two period game in Section 3.3 and solve for the infinite period game in Section 3.4. We discuss comparative static and limiting properties of the infinite period equilibrium and finally we conclude by highlighting the relation to Coase's conjecture.

3.2 The Model

In this section we first introduce a two-period model of a monopolist who sells a durable good over time and then extend it to infinite periods. Unlike the standard dynamic monopolist models, the quality of the good is not fixed but endogenously chosen by the monopolist each period. For simplicity let us assume that the marginal cost of quality $q \in \mathbb{R}^+$ is given by

$$c(q) = \frac{1}{2}q^2 \tag{3.1}$$

Hence in any period t, given the price p_t and quality q_t, the seller enjoys a mark-up

$$p_t - \frac{1}{2}q_t^2 \tag{3.2}$$

The buyers in the market are differentiated with respect to their valuation of quality. There is a continuum of buyers, where each buyer is characterized by his preference $v \in [0, 1]$. v is distributed uniformly in $[0, 1]$. If a buyer with type v buys the good at time t, she enjoys utility

$$u(v, t) = \delta^t (vq_t - p_t) \tag{3.3}$$

whereas she gets zero utility if she never purchases.

In this setting the monopolist announces a new (p_t, q_t) in each period t and the buyers who have not purchased the good until then accept or reject the offer. The monopolist does not observe the type of each buyer, but observes the set of buyers who purchase.

3.3 Two-period game

The first thing that makes the model with variable quality different than the standard dynamic monopolist model with fixed quality is that the well known "skimming property", which is the property that sorts the buyers' purchasing times with respect to their willingnesses to pay, does not automatically apply here. With a fixed quality, the skimming property would imply that the buyers with higher valuations can buy no later than buyers with low valuations due to discounting. In this model, if the quality is decreasing over time, the same argument applies. However, the seller might choose to increase the quality and in that case the buyers with high valuations might decide to wait to enjoy higher quality. We will first start this section by showing this indeed is not possible in equilibrium in the two-period model and hence the skimming property still applies.

Lemma 3.1. *In equilibrium if there are positive sales in both periods, then $\delta q_1^* \leqslant q_0^*$, where q_0^* and q_1^* indicate the first and second period qualities on the equilibrium path. In other words the discounted quality decreases.*

Proof. Suppose $\delta q_1^* > q_0^*$ and there exists a buyer with type v who buys at $t = 1$. But this means

$$v q_0^* - p_0^* \leqslant \delta \left(v q_1^* - p_1^* \right) \tag{3.4}$$

where p_0^* and p_1^* are prices on the equilibrium path. Rearranging yields

$$v \left(\delta q_1^* - q_0^* \right) \geqslant \delta p_1^* - p_0^* \tag{3.5}$$

Since $\delta q_1^* > q_0^*$, the inequality will be strictly preserved for any $v' > v$, i.e.

$$v' \left(\delta q_1^* - q_0^* \right) > \delta p_1^* - p_0^* \tag{3.6}$$

But this means all buyers with type $v' > v$ also buy at $t = 1$. This is the opposite of the typical skimming property which simply implies that high type buyers buy no earlier than low type buyers. Suppose there is a positive measure of buyers purchasing in each period. Then due to the counter skimming property, in equilibrium there exists a positive measure segment of buyers $[v_1, v_2]$ who purchase in period 0, where $0 < v_1 < v_2 < 1$. Let us analyze the seller's behavior in period 1. Due to the counter skimming property we know that the seller's equilibrium offer should be such that no buyer in $[0, v_1]$ buys in period 1. Hence the seller's offer has to exclude them by giving them a negative utility. For the high value segment $[v_2, 1]$, two things could be optimal for the seller. He either makes an offer that excludes some buyers in $[v_2, v_3]$ for some $v_3 < 1$. But this cannot be in equilibrium, because given that they are being excluded in period 1, the buyers in $[v_2, v_3]$ would rather purchase in period 0. And note that they would get a strictly positive utility in period 0, because they have a higher willingness to pay than the buyers who are already purchasing in period 0, i.e. the buyers in this case the excluded buyers, i.e. the buyers in $[v_1, v_2]$. Hence in equilibrium no buyer in $[v_2, v_3]$ can be excluded. But then constrained by covering the whole segment, the seller's optimal strategy is to extract all the surplus from the buyer with type v_2. But this means the buyer with v_2 is getting zero utility in equilibrium. Given that she is the cutoff buyer between period 0 and 1, she has to be indifferent, which means she must get zero utility if she purchases in period 0. This implies that all the buyers with $v < v_2$ get negative utility in period 0. But then in equilibrium the buyers in $[v_1, v_2]$ can not be making a purchase in period 0.

This means if $\delta q_1^* > q_0^*$, all sales has to be made in period 0 or period 1. ∎

Lemma 3.1 and its proof show that if there is a high type who waits, the seller tries to extract a high enough surplus from him that he regrets having waited. There are two degenerate cases that the lemma does not cover. One is when all the sales are made in period 0. But this clearly means zero profits for the seller and cannot be optimal. The second is where all the sales are made in period 1. This means the seller is announcing a high enough price and a low enough quality that everybody is willing to wait for period 1. However for this kind of equilibrium $\delta q_1^* > q_0^*$ is not necessary. The seller can simply announce a very high price in period 0 while offering $\delta q_1^* \leqslant q_0^*$ and still delay all the sales to period 1. Hence from this point on we can safely assume that in equilibrium $\delta q_1^* \leqslant q_0^*$, which implies that the typical skimming property holds. That means given a (p_0, q_0) pair, there exists a $v_0 \in [0,1]$ such that all buyers with $v \geqslant v_0$ buy in period 0 and buyers with $v < v_0$ wait for period 1. Let us call that marginal buyer $v_0^*(p_0, q_0)$.

We will apply backward induction starting from period 1. Suppose all buyers with $v \geqslant v_0$ buy in period 0. Then in period 1 the seller's market consists of buyers in $[0, v_0]$. If he announces a price and quality pair (p_1, q_1), the buyers with

$$v \geqslant \frac{p_1}{q_1} \tag{3.7}$$

will buy the product. Let us call the marginal buyer in period 1 $v_1^*(p_1, q_1)$.

Given v_0, the second period profits of the seller is

$$\begin{aligned} \Pi_1(p_1, q_1) &= (v_0 - v_1^*(p_1, q_1))\left(p_1 - \frac{1}{2}q_1^2\right) \\ &= \left(v_0 - \frac{p_1}{q_1}\right)\left(p_1 - \frac{1}{2}q_1^2\right) \end{aligned} \tag{3.8}$$

where the second equality is due to equation (3.7).

Proposition 3.1. *The equilibrium price and quality in period 1 are given by*

$$\begin{aligned} q_1^*(v_0) &= \frac{2}{3}v_0 \\ p_1^*(v_0) &= \frac{4}{9}v_0^2 \end{aligned}$$

where v_0 represents the marginal buyer such that all buyers with $v > v_0$ purchase in period 0.

Proof. The first order conditions with respect to p_1 and q_1 are

$$\frac{\partial \Pi_1(p_1, q_1)}{\partial p_1} = -\frac{1}{q_1}\left(p_1 - \frac{1}{2}q_1^2\right) + \left(v_0 - \frac{p_1}{q_1}\right) = 0 \tag{3.9}$$

$$\frac{\partial \Pi_1(p_1, q_1)}{\partial q_1} = \frac{p_1}{q_1}\left(p_1 - \frac{1}{2}q_1^2\right) - q_1\left(v_0 - \frac{p_1}{q_1}\right) = 0 \tag{3.10}$$

From the first FOC we can derive that at the optimum

$$p_1(q_1) = \frac{q_1^2}{4} + \frac{v_0 q_1}{2} \tag{3.11}$$

Rearranging and dividing the two FOCs side by side gives

$$p_1 = q_1^2 \tag{3.12}$$

Plugging this back in the previous equation we get at the optimum q_1^*

$$p_1(q_1^*) = q_1^{*2} = \frac{q_1^{*2}}{4} + \frac{v_0 q_1^*}{2} \tag{3.13}$$

Solving this polynomial and also checking the second order condition yields

$$q_1^* (v_0) = \frac{2v_0}{3} \tag{3.14}$$

$$p_1^* (v_0) = \frac{4v_0^2}{9} \tag{3.15}$$

which are the expressions given in the proposition. ∎

Not surprisingly both the second period quality and price increase as there are more high value buyers left. The second period profit which is given by $\frac{2v_0^3}{27}$ also increases in v_0.

Now we can move back to period 0. Recall that given the skimming property, $v_0^* (p_0, q_0)$ represents the marginal buyer. Given an announcement of (p_0, q_0), $v_0^* (p_0, q_0)$ can be found by the following indifference condition:

$$v_0^* q_0 - p_0 = \delta \left(v_0^* q_1^* (v_0^*) - p_1^* (v_0^*) \right) \tag{3.16}$$

When we plug in the equilibrium values of $q_1^* (v_0^*)$ and $p_1^* (v_0^*)$, this indifference condition generates two possible equilibrium values for v_0^*, but one of them can be shown to violate Lemma 3.1. This reduces the marginal buyer to be

$$v_0^* (p_0, q_0) = \frac{9q_0 - 3\sqrt{9q_0^2 - 8\delta p_0}}{4\delta} \tag{3.17}$$

We can directly observe that the type of the marginal buyer is increasing in price and decreasing in quality, i.e. as the offer in period 0 gets less desirable, more buyers wait for period 1.

Now we can write the two-period profits as a function of (p_0, q_0):

$$\begin{aligned}\Pi_0(p_0,q_0) &= (1-v_0^*(p_0,q_0))\left(p_0-\frac{1}{2}q_0^2\right) \\ &\quad +\delta\left(v_0^*(p_0,q_0)-v_1^*(p_1,q_1)\right)\left(p_1^*(v_0^*(p_0,q_0))-\frac{1}{2}q_1^*(v_0^*(p_0,q_0))\right) \\ &= (1-v_0^*(p_0,q_0))\left(p_0-\frac{1}{2}q_0^2\right)+\frac{2\delta}{27}\left(v_0^*(p_0,q_0)\right)^3\end{aligned} \tag{3.18}$$

Proposition 3.2. *The equilibrium price and quality in period 0 are given by*

$$\begin{aligned}q_0^*(\delta) &= \frac{2(9-4\delta)}{27-16\delta} \\ p_0^*(\delta) &= \frac{4(9-2\delta)(9-4\delta)^2}{9(27-16\delta)^2}\end{aligned}$$

Proof. Taking the first order conditions with respect to (p_0, q_0) gives

$$\frac{\partial\Pi_0(p_0,q_0)}{\partial p_0} = \left(\begin{array}{c}(1-v_0^*(p_0,q_0)) \\ -\left(p_0-\frac{1}{2}q_0^2-\frac{2\delta}{9}\left(v_0^*(p_0,q_0)\right)^2\right)\frac{\partial\Pi_0 v_0^*(p_0,q_0)}{\partial p_0}\end{array}\right)=0 \tag{3.19}$$

$$\frac{\partial\Pi_0(p_0,q_0)}{\partial q_0} = \left(\begin{array}{c}-q_0(1-v_0^*(p_0,q_0)) \\ -\left(p_0-\frac{1}{2}q_0^2-\frac{2\delta}{9}\left(v_0^*(p_0,q_0)\right)^2\right)\frac{\partial\Pi_0 v_0^*(p_0,q_0)}{\partial q_0}\end{array}\right)=0 \tag{3.20}$$

where $v_0^*(p_0, q_0)$ is as defined in (3.17). Rearranging the two FOCs yield

$$\frac{\partial v_0^*(p_0^*,q_0^*)}{\partial q_0} = -q_0\frac{\partial v_0^*(p_0^*,q_0^*)}{\partial p_0} \tag{3.21}$$

Applying this to (3.17) we get

$$p_0^* = \left(1-\frac{2}{9}\delta\right)q_0^{*2} \tag{3.22}$$

which can together with (3.17) and (3.21) be used to show

$$v_0^*\left(p_0^*, q_0^*\right) = q_0^* \tag{3.23}$$

Finally using these equations it can be found that

$$q_0^*\left(\delta\right) = \frac{2\left(9-4\delta\right)}{27-16\delta} \tag{3.24}$$

$$p_0^*\left(\delta\right) = \frac{4\left(9-2\delta\right)\left(9-4\delta\right)^2}{9\left(27-16\delta\right)^2} \tag{3.25}$$

which are the expressions given in the proposition. ∎

Using the expressions given in Proposition 3.2, we can deduce that

$$\frac{\partial q_0^*\left(\delta\right)}{\partial \delta} > 0 \tag{3.26}$$

$$\frac{\partial p_0^*\left(\delta\right)}{\partial \delta} > 0 \tag{3.27}$$

This means that both the quality and the price in period 0 increase in the patience level. The indirect profits can be derived as

$$\frac{2}{27}\frac{\left(9-4\delta\right)^3}{\left(27-16\delta\right)^2} \tag{3.28}$$

which first decreases and then increases in δ. This dynamic with respect to δ is no different than the dynamic we would see in a two-period fixed quality model. In the fixed quality model at $\delta = 0$, the seller gets the static monopoly profit, which is what he would achieve if he could commit. What is interesting is that as $\delta \to 1$, the seller's payoff approaches to the commitment profit again. The reason is that at the limit, at $\delta = 1$, the seller announces a high enough price in period 0 so that all the buyers wait till period 1. Then in period 1 he gets his commitment profit, i.e. the static monopoly profit.

In the case of endogenous quality, at $\delta = 0$, the seller starts with the static monopoly profit just like the fixed quality case. However, the static monopoly profit is no longer necessarily the monopolist's commitment profit. Because due to the "adverse selection" element inherent in the problem, the seller can use the two time periods to separate out buyers with higher willingness to pay and hence extract more surplus. For example when $\delta = 1$, the seller's commitment quality and price scheme is a menu of two contracts (two price-quality couples) distributed over time. So as $\delta \to 1$, the seller's profits are strictly above that of a static monopolist who is restricted to offer only one contract.

3.3.1 Discussion of the T-period Game

Just like in the two-period model, the alternative of delaying all the sales until the last period exists for any finite number of periods and hence the convexity of profits in δ prevails. Although we do not explicitly solve for the general T-period game, we would like to discuss structure of the equilibrium. One can recursively apply backward induction beyond two periods, but this implicitly assumes that the skimming property holds for any number of periods. Therefore, we first assure that this property holds by showing that the discounted quality decreases on the equilibrium path.

Lemma 3.2. *In the T-period game, in equilibrium, if there are positive sales in periods t and t', where $t < t'$, then $\delta^t q_t^* > \delta^{t'} q_{t'}^*$, where q_t^* and $q_{t'}^*$ indicate the qualities on the equilibrium path.*

Proof. We prove this in 4 steps.

Step 1: There cannot be disjoint segments of buyers who remain in the market at T.

To see this let $[\overline{v}, \overline{v}']$ represent the segment of highest type buyers who still have not purchased by period T. Suppose there exists another segment disjoint from the highest segment and let $[\underline{\beta}, \underline{\beta}']$ be the second highest segment that still has not bought, where $0 \leqslant \underline{\beta} < \underline{\beta}' < \overline{v}$. This means by definition there exists a buyer with type $\widehat{v}$ who purchased at a time $t < T$, where $\underline{\beta}' < \widehat{v} < \overline{v}'$.

Take buyer $\widehat{v}$, the fact that she purchased at $t < T$ implies that

$$\widehat{v}\left(\delta^t q_t^* - \delta^T q_T^*\right) \geqslant \delta^t p_1^* - \delta^T p_0^* \tag{3.29}$$

Suppose $\delta^t q_t^* - \delta^T q_T^* > 0$. Then by (3.29) all buyers with $v > \widehat{v}$, prefer buying at t to T. But this contradicts with the fact that buyers in $[\overline{v}, \overline{v}']$ have waited for T.

Hence it has to be true that $\delta^t q_t^* - \delta^T q_T^* \leqslant 0$. Now we will show that the buyers in $[\underline{\beta}, \underline{\beta}']$ must be excluded at T. To see this suppose otherwise. The fact that the buyers in $[\underline{\beta}, \underline{\beta}']$ are not excluded means that at T they get a positive payoff. But due to the fact that $\delta^t q_t^* - \delta^T q_T^* \leqslant 0$, we also know by (3.29) that they prefer buying at period t to T. Hence they must be getting a positive payoff at t too. But this contradicts the fact that buyers in $[\underline{\beta}, \underline{\beta}']$ have waited for T.

Now we know that if there are disjoint segments left by T, then $\delta^t q_t^* - \delta^T q_T^* \leqslant 0$ and only the highest segment can be served at T. Given segment $[\overline{v}, \overline{v}']$, the seller will either find it optimal to exclude some lower types in $[\overline{v}, \overline{v}']$ or serve them all by excluding all the surplus from $\overline{v}$.In both cases there exists at least one buyer who ends up with zero surplus. Given the fact that there exists a buyer with type $\widehat{v} < \overline{v}$ who purchased at t, this zero surplus buyer could have made a positive surplus at t. But this contradicts the fact that she has waited.

Now we know that there cannot be disjoint sets of buyers who have not purchased before T.

Step 2: The set buyers who remain in the market at T is a segment of the form $[0, v]$ for some $v \leqslant 1$.

We know by Step 1 that the buyers who have not purchased before T has to be a segment. Suppose it is a segment of the form $[v, v']$ where $0 < v < v' \leqslant 1$. But then once again seller's optimal offer in T will end some buyer with zero surplus, who hence regrets having waited given that there are buyers with strictly lower types who bought earlier. So this cannot be an equilibrium.

Step 3: There cannot be disjoint segment of buyers who remain in the market at $T-1$.

Let $[0, v]$ be the set of buyers who wait till period T on the equilibrium path. Then all the gaps in $[v, 1]$, if there are any, has to be closed in period $T-1$. Recall that in step 1 the first argument that relied on (3.29) implied that the lower segments should be excluded. Hence the seller cannot serve to two disjoint segments in period $T-1$. But this means the seller arrives at period T with disjoint segments which is a contradiction. Suppose there exists only one segment $[\widehat{v}, \widehat{v}']$ other than $[0, v]$ that has not purchased before $T-1$, where $v < \widehat{v} < \widehat{v}' \leqslant 1$. Optimality of the seller's offer implies that if the seller covers the whole gap, he will due so by extracting enough such that either $\widehat{v}$ or $\widehat{v}'$ is indifferent between buying at $T-1$ and at T. Now suppose the indifferent buyer is $\widehat{v}$. This implies

$$\widehat{v}\left(\delta^{T-1} q_{T-1}^{*} - \delta^{T} q_{T}^{*}\right) = \delta^{T-1} p_{T-1}^{*} - \delta^{T} p_{T}^{*} \tag{3.30}$$

Notice that there exists a segment of buyers $(\widetilde{v}, \widehat{v})$ with $\widetilde{v} < \widehat{v}$ who purchased at time $t < T-1$. This implies

$$\widehat{v}\left(\delta^{T-1} q_{T-1}^{*} - \delta^{t} q_{t}^{*}\right) = \delta^{T-1} p_{T-1}^{*} - \delta^{t} p_{t}^{*} \tag{3.31}$$

Combining this with the previous equality for $\widehat{v}$ yields

$$\widehat{v}\left(\delta^{T} q_{T}^{*} - \delta^{t} q_{t}^{*}\right) = \delta^{T} p_{T}^{*} - \delta^{t} p_{t}^{*} \tag{3.32}$$

Depending on the sign of $\left(\delta^{T} q_{T}^{*} - \delta^{t} q_{t}^{*}\right)$, all buyers in $[0, \widehat{v}]$ will either prefer T or t. But this contradicts the fact that $(\widetilde{v}, \widehat{v})$ buy at t and some buyers in $[0, v]$ buy at T.

Suppose the indifferent buyer is $\widehat{v}'$ this implies

$$\widehat{v}'\left(\delta^{T-1} q_{T-1}^{*} - \delta^{T} q_{T}^{*}\right) = \delta^{T-1} p_{T-1}^{*} - \delta^{T} p_{T}^{*} \tag{3.33}$$

Depending on the sign of $\delta^{T-1} q_{T-1}^{*} - \delta^{T} q_{T}^{*}$, all the buyers in $[0, \widehat{v}']$ will prefer either T or $T-1$. This contradicts the fact that $[\widehat{v}, \widehat{v}']$ buy at $T-1$ and some buyers in $[0, v]$ buy at T.

Step 4: The argument for Step 3 applies to all periods $t \leqslant T-1$. Hence in equilibrium, there exists a decreasing sequence $\{v_t\}$ such that the buyers who have purchased before time t can be represented by the segment $[v_t, 1]$. Take two periods $t < t'$ such that there are positive sales in both periods but zero sales in any period between those periods. But if this is true the optimality of the seller's offers at each t imply that buyer $v_{t'}$ should be indifferent between buying at t and t' which can be represented as

$$v_t\left(\delta^{t} q_{t}^{*} - \delta^{t'} q_{t'}^{*}\right) = \delta^{t} p_{t}^{*} - \delta^{t'} p_{t'}^{*} \tag{3.34}$$

Since v_t is a decreasing sequence we also know that all buyers with higher types than v strictly prefer t to $t+1$ and all buyers with lower types prefer $t+1$. But given the indifference condition of v_{t+1} this can only be true if

$$\delta^t q_t^* - \delta^{t\prime} q_{t\prime}^* > 0 \tag{3.35}$$

which proves the lemma for any two periods of consecutive sales. For all periods there exists another period of consecutive sales and thus all the periods could be ranked with respect to the discounted quality as given in the lemma. ∎

Having proved the skimming property, recursive backward induction of the game beyond two periods would reveal first that the subgame perfect equilibrium is unique (due to the uniqueness of equilibrium in period T) and second that the equilibrium price and quality offers at period t have following functional forms:

$$q_t(v_t) = y_t v_t \tag{3.36}$$

$$p_t(v_t) = x_t v_t^2 \tag{3.37}$$

where y_t and x_t are functions of t, which constitute solutions to some optimality conditions.

We would like to emphasize the skimming property and the structure of the finite period equilibrium because in the next section we will discuss the infinite-period equilibrium which constitutes the limit to the finite-period equilibrium.

3.4 Infinite-period game

In this section we analyze the infinite game. Although marginal cost of the good is not exogenously given, the variable quality model that we investigate here has the flavor of the "no-gap case" discussed in Gul *et al.* *(1986)*. No matter what quality is chosen, within an upper bound, there will always be a buyer type who has the exact valuation as the marginal cost. This implies that the market might stay open forever. Hence there are possibly multiple equilibria. Gul *et al.* *(1986)* look into a specific class of stationary equilibria and show that the Coase conjecture applies to that specific class. Here we take a simpler approach. We concentrate on one stationary equilibrium, which we then prove constitutes a limit to the finite-period equilibrium.

Given that we are looking for a limiting equilibrium, we would like to preserve the skimming property that we proved in the previous section. Hence we look for a stationary equilibrium of this game with the property that in each period t, buyers with type $v \geq v_{t+1}$ purchase if they have not already purchased, where v_{t+1} is a function that is decreasing in t. In other words v_{t+1} represents the market penetration achieved by time t. Given this definition of v_t, a stationary strategy for the monopolist is $(p_t, q_t) : [0,1] \to \mathbb{R}_+^2$ and a strategy for a buyer with type v consists of an accept-reject rule for each price quality at each time t, i.e. $\mathbb{R}_+^2 \to \{0,1\}$ pair. In other words the stationary strategies can be summarized by the following rules:

1. $p_t(v_t)$

2. $q_t(v_t)$

3. $v_{t+1}(p_t, q_t)$

where the last one is the market penetration achieved by time t, which is deduced from the accept-reject rules of the buyers. From this point on whenever we write p_t, q_t and v_{t+1}, we refer to these functions.

Given the stationary strategies, monopolist's infinite-period profits starting at period t can be represented by the following value function:

$$\Pi_t(v_t) = \max_{p_t, q_t} \left\{ (v_t - v_{t+1}) \left(p_t - \frac{1}{2} q_t^2 \right) + \delta \Pi_{t+1}(v_{t+1}) \right\} \tag{3.38}$$

where v_t is the state variable and p_t and q_t are the choice variables.

Let us first look at the first order conditions of the objective with respect to p_t and q_t respectively

$$\begin{aligned} -\frac{\partial v_{t+1}}{\partial p_t} \left(p_t - \frac{1}{2} q_t^2 \right) + (v_t - v_{t+1}) + \delta \frac{\partial \Pi_{t+1}}{\partial v_{t+1}} \frac{\partial v_{t+1}}{\partial p_t} &= 0 \\ -\frac{\partial v_{t+1}}{\partial q_t} \left(p_t - \frac{1}{2} q_t^2 \right) - q_t (v_t - v_{t+1}) + \delta \frac{\partial \Pi_{t+1}}{\partial v_{t+1}} \frac{\partial v_{t+1}}{\partial q_t} &= 0 \end{aligned} \tag{3.39}$$

We can then use the envelope theorem to get

$$\frac{\partial \Pi_t(v_t)}{\partial v_t} = p_t - \frac{1}{2} q_t^2 \tag{3.40}$$

Finally we know that v_{t+1} should be indifferent between t and $t+1$, which can be summarized as

$$v_{t+1} q_t - p_t = \delta \left(v_{t+1} q_{t+1} - p_{t+1} \right) \tag{3.41}$$

Since we are looking for an equilibrium with v_{t+1} decreasing, the following transversality condition must hold:

$$\lim_{t\to\infty} v_{t+1}\left(p_t\left(v_t\right), q_t\left(v_t\right)\right) = 0 \tag{3.42}$$

Using the envelope theorem to substitute for $\frac{\partial \Pi_{t+1}}{\partial v_{t+1}}$, the equilibrium can be summarized by the following four equations in p_t, q_t and v_{t+1}:

$$(v_t - v_{t+1}) - \left(\left(p_t - \frac{1}{2}q_t^2\right) - \delta\left(p_{t+1} - \frac{1}{2}q_{t+1}^2\right)\right)\frac{\partial v_{t+1}}{\partial p_t} = 0 \tag{3.43}$$

$$-q_t\left(v_t - v_{t+1}\right) - \left(\left(p_t - \frac{1}{2}q_t^2\right) - \delta\left(p_{t+1} - \frac{1}{2}q_{t+1}^2\right)\right)\frac{\partial v_{t+1}}{\partial q_t} = 0 \tag{3.44}$$

$$v_{t+1}q_t - p_t = \delta\left(v_{t+1}q_{t+1} - p_{t+1}\right) \tag{3.45}$$

$$\lim_{t\to\infty} v_{t+1}\left(p_t\left(v_t\right), q_t\left(v_t\right)\right) = 0 \tag{3.46}$$

which are the two FOC, the indifference condition and the transversality condition.

Lemma 3.3. *The following hold for the equilibrium rules:*

1. $\frac{\partial v_{t+1}}{\partial q_t} = -q_t \frac{\partial v_{t+1}}{\partial p_t}$

2. $v_{t+1}\left(p_t\left(v_t\right), q_t\left(v_t\right)\right) = q_t\left(v_t\right)$

Proof. Point 1 follows directly from the rearrangement of the two FOCs, equations (3.43) and (3.44).

To see that point 2 holds let us use the implicit function theorem to derive $\frac{\partial v_{t+1}}{\partial q_t}$ and $\frac{\partial v_{t+1}}{\partial p_t}$ through equation (3.45). Let

$$f\left(\delta, v_{t+1}, q_{t+1}, p_{t+1}, q_t, p_t\right) = v_{t+1}q_t - p_t - \delta\left(v_{t+1}q_{t+1} - p_{t+1}\right) \tag{3.47}$$

$$\begin{aligned}\frac{\partial v_{t+1}}{\partial q_t} &= -\frac{\frac{\partial f(\delta,v_{t+1},q_{t+1},p_{t+1},q_t,p_t)}{\partial q_t}}{\frac{\partial f(\delta,v_{t+1},q_{t+1},p_{t+1},q_t,p_t)}{\partial v_{t+1}}} \\ &= -\frac{v_{t+1}}{q_t - \delta\left(q_{t+1} + v_{t+1}\frac{dq_{t+1}}{dv_{t+1}} - \frac{dp_{t+1}}{dv_{t+1}}\right)}\end{aligned} \tag{3.48}$$

and similarly

$$\begin{aligned}\frac{\partial v_{t+1}}{\partial p_t} &= -\frac{\frac{\partial f(\delta,v_{t+1},q_{t+1},p_{t+1},q_t,p_t)}{\partial p_t}}{\frac{\partial f(\delta,v_{t+1},q_{t+1},p_{t+1},q_t,p_t)}{\partial v_{t+1}}} \\ &= -\frac{-1}{q_t - \delta\left(q_{t+1} + v_{t+1}\frac{dq_{t+1}}{dv_{t+1}} - \frac{dp_{t+1}}{dv_{t+1}}\right)}\end{aligned} \tag{3.49}$$

Dividing (3.48) by (3.49) we can derive that

$$\frac{\partial v_{t+1}}{\partial q_t} = -v_{t+1}\frac{\partial v_{t+1}}{\partial p_t} \tag{3.50}$$

But we know that equilibrium rules $p_t(v_t)$ and $q_t(v_t)$ satisfy point 1 of the lemma. That together with (3.50) implies

$$v_{t+1}\left(p_t(v_t), q_t(v_t)\right) = q_t(v_t) \tag{3.51}$$

which proves the lemma. ∎

As we argued above, the equilibrium rules have to satisfy the four equations (3.43)-(3.46). The properties given in the lemma are useful because the first property can replace one of the two FOCs. But then we could further replace it with the second property given that equation (3.45) has to be satisfied. Hence we can say that if a set of strategies satisfy (3.44)-(3.46) and point (2) in Lemma 3.3, they constitute an equilibrium.

Proposition 3.3. *The following rules constitute a stationary equilibrium:*

1. $p_t^*(v_t) = x(\delta) v_t^2$

2. $q_t^*(v_t) = y(\delta) v_t$

3. $v_{t+1}^*(p_t, q_t) = \frac{q_t - \sqrt{q_t^2 - 4\delta(y(\delta)-x(\delta))p_t}}{2\delta(y(\delta)-x(\delta))}$ *where*

$$x(\delta) = \frac{y(\delta)^2 (1 - \delta y(\delta))}{1 - \delta y(\delta)^2}$$

and $y(\delta)$ *is the unique solution to*

$$\delta y(\delta)^3 - 3y(\delta) + 2 = 0$$

$$0 \leqslant y(\delta) \leqslant 1$$

Proof. Let us start with point (3). Given the equilibrium pricing and quality rules the indifference condition in equation (3.45) can be written as

$$v_{t+1}q_t - p_t = \delta \left(v_{t+1}y(\delta) v_{t+1} - x(\delta) v_{t+1}^2\right) \tag{3.52}$$

First notice that the individual rationality for buyer with type v_{t+1} implies

$$(y(\delta) - x(\delta)) \geqslant 0$$

The indifference condition can be rewritten as the following polynomial in v_{t+1}:

$$-\delta (y(\delta) - x(\delta)) v_{t+1}^2 + q_t v_{t+1} - p_t = 0 \tag{3.53}$$

which has the following two roots:

$$v_{t+1} = \frac{q_t \pm \sqrt{q_t^2 - 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)p_t}}{2\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)} \tag{3.54}$$

Clearly if $q_t^2 - 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)p_t < 0$, there exists no real solution, which means that all the buyers strictly prefer period $t+1$ to t. In other words there would be no sales in period t. But this cannot be optimal for the seller, because he can always announce p_{t+1}, q_{t+1} at period t and increase his profits by a factor of $\frac{1}{\delta}$ since the strategies are stationary. Hence we can safely assume $q_t^2 - 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)p_t \geqslant 0$.

To see which root to pick we will refer to the transversality condition in equation (3.46). If we plug in the rules $p_t\left(v_t\right)$ and $q_t\left(v_t\right)$ into equation (3.54) we get the following deduced form for v_{t+1}:

$$v_{t+1}\left(p_t\left(v_t\right), q_t\left(v_t\right)\right) = \frac{y\left(\delta\right) \pm \sqrt{y\left(\delta\right)^2 - 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)x\left(\delta\right)}}{2\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)} v_t \tag{3.55}$$

But then equation (3.46) is equivalent to the requirement that

$$0 \leqslant \frac{y\left(\delta\right) \pm \sqrt{y\left(\delta\right)^2 - 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)x\left(\delta\right)}}{2\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)} < 1 \tag{3.56}$$

We know that $\left(y\left(\delta\right) - x\left(\delta\right)\right) \geqslant 0$. Let us look at the first root.

$$\frac{y\left(\delta\right) + \sqrt{y\left(\delta\right)^2 - 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)x\left(\delta\right)}}{2\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)} - 1$$
$$= \frac{y\left(\delta\right) + \sqrt{y\left(\delta\right)^2 - 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)x\left(\delta\right)} - 2\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)}{2\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)}$$

If $y\left(\delta\right) > 2\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)$, then the difference is positive and the transversality is violated, so suppose $y(\delta) \leqslant 2\delta(y(\delta) - x(\delta))$. Then we can look at the differences of squares again, which gives

$$y(\delta)^2 - 4\delta(y(\delta) - x(\delta))x(\delta) - (2\delta(y(\delta) - y(\delta)) - y(\delta))^2$$
$$= -4\delta(y(\delta) - x(\delta))x(\delta) - 4\delta^2(y(\delta) - x(\delta))^2 + 4\delta(y(\delta) - x(\delta))y(\delta)$$
$$= 4\delta(y(\delta) - x(\delta))^2(1 - \delta) > 0$$

which again implies that the root is larger than unity. Now let us look at the second root

$$\frac{y(\delta) - \sqrt{y(\delta)^2 - 4\delta(y(\delta) - x(\delta))x(\delta)}}{2\delta(y(\delta) - x(\delta))} - 1$$
$$= \frac{y(\delta) - \sqrt{y(\delta)^2 - 4\delta(y(\delta) - x(\delta))x(\delta)} - 2\delta(y(\delta) - x(\delta))}{2\delta(y(\delta) - x(\delta))}$$

If $y(\delta) < 2\delta(y(\delta) - x(\delta))$, the difference is negative so transversality is satisfied. If $y(\delta) \geqslant 2\delta(y(\delta) - x(\delta))$, the difference of squares is

$$4\delta^2(y(\delta) - x(\delta))^2 - 4\delta(y(\delta) - x(\delta))y(\delta) + 4\delta(y(\delta) - x(\delta))x(\delta)$$
$$= -4\delta(y(\delta) - x(\delta))^2(1 - \delta) < 0$$

Hence transversality is satisfied. Hence we can say that the rules given in the proposition satisfy both the indifference and transversality conditions given in equations (3.45) and (3.46) iff

$$(y(\delta) - x(\delta)) \geqslant 0 \tag{3.57}$$

Now we need to check whether they satisfy the FOCs. As we argued above, given (3.45), checking whether equations (3.43) and (3.44) hold is equivalent to checking for equations (3.44) and point (2) of Lemma 3.3. Let us first look at that point of Lemma 3.3. Notice that

$$v_{t+1}\left(p_t\left(v_t\right), q_t\left(v_t\right)\right) = \frac{y\left(\delta\right) - \sqrt{y\left(\delta\right)^2 - 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)x\left(\delta\right)}}{2\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)} v_t \tag{3.58}$$

But then Lemma 3.3 implies that

$$\frac{y\left(\delta\right) - \sqrt{y\left(\delta\right)^2 - 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)x\left(\delta\right)}}{2\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)} = y\left(\delta\right) \tag{3.59}$$

First of all this clearly requires that

$$\left(1 - 2\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)\right) > 0 \tag{3.60}$$

Rearranging the terms we get

$$y\left(\delta\right)\left(1 - 2\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)\right) - \sqrt{y\left(\delta\right)^2 - 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)x\left(\delta\right)} = 0 \tag{3.61}$$

Squaring both terms

$$\begin{aligned} y\left(\delta\right)^2\left(1 - 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right) + 4\delta^2\left(y\left(\delta\right) - x\left(\delta\right)\right)^2\right) - y\left(\delta\right)^2 + 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)x\left(\delta\right) & & (3.62) \\ 4\delta\left(y\left(\delta\right) - x\left(\delta\right)\right)\left(\delta y\left(\delta\right)^3 - \delta y\left(\delta\right)^2 x\left(\delta\right) - y\left(\delta\right)^2 + x\right) &= 0 \end{aligned}$$

We know that $\left(y\left(\delta\right) - x\left(\delta\right)\right) > 0$ has to hold otherwise at period 0 no buyer would purchase because they would get a negative utility. But given that the equilibrium is stationary this means no buyer would ever purchase generating zero profits for the seller. Thus it has to be the case that

$$\delta y(\delta)^3 - \delta y(\delta)^2 x(\delta) - y(\delta)^2 + x = 0 \tag{3.63}$$

which automatically implies that

$$x(\delta) = \frac{y^2(1 - \delta y(\delta))}{1 - \delta y(\delta)^2} \tag{3.64}$$

Having $x(\delta)$ defined, the condition in (3.57) can be reduced to

$$0 \leqslant y(\delta) \leqslant 1 \tag{3.65}$$

Now we have to check for the second condition of optimality, namely equation (3.44). Let us plug in $p_t(v_t)$, $q_t(v_t)$ and $v_{t+1}(p_t(v), q_t(v))$ in equation (3.44) to get

$$\begin{aligned} -q_t(v_t - v_{t+1}) - \left(\left(p_t - \frac{1}{2}q_t^2\right) - \delta\left(p_{t+1} - \frac{1}{2}q_{t+1}^2\right)\right)\frac{\partial v_{t+1}}{\partial q_t} &= 0 \quad (3.66) \\ -y(\delta)(1 - y(\delta)) - \left(\begin{array}{c} \left(x(\delta) - \frac{1}{2}y(\delta)^2 - \delta\left(x(\delta)y(\delta)^2 - \frac{1}{2}y(\delta)^4\right)\right) \\ \left(\frac{\sqrt{y(\delta)^2 - 4\delta(y(\delta) - x(\delta))x(\delta)} - y}{2\delta(y(\delta) - x(\delta))\sqrt{y(\delta)^2 - 4\delta(y(\delta) - x(\delta))x(\delta)}}\right) \end{array}\right) &= 0 \end{aligned}$$

Substituting for $x(\delta)$ reduces this expression to

$$y\left(1 + \delta y^2 - 2\delta y\right)\frac{2 - 3y + \delta y^3}{-1 + \delta y^2} = 0 \tag{3.67}$$

The term on the left is always positive since $r < 1$. Hence the condition reduces to

$$\delta y^3 - 3y + 2 = 0 \tag{3.68}$$

We know that the condition in (3.65) has to be satisfied. We will show that there exists a unique real root that satisfies both (3.65) and (3.68). Notice that the polynomial in (3.68) gets value 2 for $y = 0$ and $\delta - 1 \leqslant 0$ for $y = 1$. But this means there exists a $0 \leqslant y(\delta) \leqslant 1$ such that (3.68) is satisfied. Now let us look at how the polynomial changes in y between 0 and 1. The derivative of the polynomial with respect to y is $3(\delta y^2 - 1) < 0$, which implies that the function is decreasing between 0 and 1. Hence $y(\delta)$ is the unique solution to the optimization. ∎

Notice that this equilibrium has the same functional form as the equilibrium in the finite period game. Before going into the implications and the properties of this equilibrium we show that it indeed is the limiting equilibrium for the finite-period game.

Proposition 3.4. *Let $p_t(T, v_t), q_t(T, v_t)$ and $v_{t+1}(T, p_t, q_t)$ represent an equilibrium of the T-period game. Then for all $t \leqslant T$,*

$$\begin{aligned}\lim_{T\to\infty} p_t(T, v_t) &= p_t^*(v_t)\\ \lim_{T\to\infty} q_t(T, v_t) &= q_t^*(v_t)\\ \lim_{T\to\infty} v_{t+1}(T, p_t, q_t) &= v_{t+1}^*(p_t, q_t)\end{aligned}$$

where $p_t^*(v_t)$, $q_t^*(v_t)$ and $v_{t+1}^*(p_t, q_t)$ are as defined in Proposition 3.3.

Proof. First of all, by Lemma 3.2 we know that $v_{t+1}(T, p_t, q_t)$ must be decreasing. And by backward induction we know that the equilibrium is unique and has the form

$$q_t(T, v_t) = y_t(T) v_t \tag{3.69}$$

$$p_t(T, v_t) = x_t(T) v_t^2 \tag{3.70}$$

The equilibrium for the $T-$period game then can be characterized by $\{x_t(T), y_t(T)\}$, which are two sequences in t, and a market penetration function $v_{t+1}(T, p_t, q_t)$, which will be given by the indifference condition. Notice that the value function, first order conditions and the indifference condition in equations (3.43)-(3.45)apply to the T-period game as well. Hence the conditions in Lemma 3.3 also apply, which means

$$v_{t+1}(T, p_t(T, v_t), q_t(T, v_t)) = y_t(T) v_t \tag{3.71}$$

But the skimming property we proved in Lemma 3.2 imply that $y_t(T) < 1$ for all t and T. Now let us look at the indifference condition which identifies buyer $v_{t+1}(T, p_t, q_t)$

$$v_{t+1} q_t - p_t = \delta v_{t+1}^2 (y_{t+1}(T) - x_{t+1}(T)) \tag{3.72}$$

which by the individual rationality constraint implies that $y_{t+1}(T) \geqslant x_{t+1}(T)$ for all t. From the indifference condition if we solve for v_{t+1}, we get

$$v_{t+1} = \frac{q_t \pm \sqrt{q_t^2 - 4\delta (y_{t+1}(T) - x_{t+1}(T)) p_t}}{2\delta (y_{t+1}(T) - x_{t+1}(T))} \tag{3.73}$$

However, the additive root cannot be part of the equilibrium because when we plug in the equilibrium values for p_t and q_t we get

$$\frac{y_t(T)+\sqrt{y_t^2(T)-4\delta\left(y_{t+1}(T)-x_{t+1}(T)\right)x_t(T)}}{2\delta\left(y_{t+1}(T)-x_{t+1}(T)\right)}v_t \tag{3.74}$$

However since $y_t(T)<1$, $\frac{y_t(T)+\sqrt{y_t^2(T)-4\delta(y_{t+1}(T)-x_{t+1}(T))x_t(T)}}{2\delta(y_{t+1}(T)-x_{t+1}(T))}>1$, which violates the skimming property. Hence the equilibrium cut-off rule is given by

$$v_{t+1}(T,p_t,q_t)=\frac{q_t-\sqrt{q_t^2-4\delta\left(y_{t+1}(T)-x_{t+1}(T)\right)p_t}}{2\delta\left(y_{t+1}(T)-x_{t+1}(T)\right)} \tag{3.75}$$

which again is the same as what we had in the infinite equilibrium except for the t indices for x and y functions. We know that

$$v_{t+1}\left(T,p_t\left(T,v_t\right),q_t\left(T,v_t\right)\right)=y_t(T)\,v_t \tag{3.76}$$

This with the pervious equation implies

$$x_t(T)-\left(1-\delta\left(y_{t+1}(T)-x_{t+1}(T)\right)\right)y_t^2(T)=0 \tag{3.77}$$

Plugging the equilibrium price and quality functions and the-cutoff rule in the FOC given by (3.44) we get the following

$$\left(2x_t(T)-y_t^2(T)\right)\left(2-3y_t(T)\right)+\delta y_t^3(T)\left(2x_{t+1}(T)-y_{t+1}^2(T)\right)=0 \tag{3.78}$$

Hence the T-period equilibrium is characterized by the solution to the system of difference equations given in (3.77) and (3.78) plus the terminal conditions that

$$x_T(T) = \frac{4}{9} \tag{3.79}$$

$$y_T(T) = \frac{2}{3} \tag{3.80}$$

Now that we know something about the structure of the equilibrium we can discuss the limit. First of all let $\Pi_t(T, v)$ represent the equilibrium value function at time t given market penetration v in the T-period game. It is clear that $\Pi_t(T, v)$ is continuous in T and also that due to discounting $\lim_{T\to\infty} \Pi_t(T, v)$ is well defined. This means that

$$\lim_{T\to\infty} \Pi_t(T, v) = \Pi_t\left(\lim_{T\to\infty} T, v\right) = \Pi_t\left(\lim_{T\to\infty} T - 1, v\right) = \lim_{T\to\infty} \Pi_t(T-1, v) \tag{3.81}$$

where the second and last equalities are due to continuity and the third one is due to $\lim_{T\to\infty} T = \lim_{T\to\infty} T - 1$.

Now let $\widehat{\Pi}_t(T, v, p)$ represent the maximized value function over q given p and v. Then

$$\lim_{T\to\infty} \widehat{\Pi}_t(T, v, p(T-1, v)) = \lim_{T\to\infty} \widehat{\Pi}_t(T-1, v, p(T-1, v)) = \lim_{T\to\infty} \Pi_t(T-1, v) \tag{3.82}$$

where the third equality is due to the fact that by definition

$$\widehat{\Pi}_t(T-1, v, p(T-1, v)) = \Pi_t(T-1, v) \tag{3.83}$$

But then by the squeezing theorem it has to be the case that

$$\lim_{T\to\infty} \Pi_t(T, v) = \lim_{T\to\infty} \widehat{\Pi}_t(T, v, p(T-1, v)) \tag{3.84}$$

However notice that we can also write $\Pi_t(T,v)$ as $\Pi_t(T,v,p(T,v))$ which means

$$\lim_{T\to\infty}\widehat{\Pi}_t(T,v,p(T,v))=\lim_{T\to\infty}\widehat{\Pi}_t(T,v,p(T-1,v)) \tag{3.85}$$

However, by definition $\widehat{\Pi}_t(T,v,p(T,v))$ is the unique global maximum of the function $\widehat{\Pi}_t$. But then for the previous limit equality to hold it has to be the case that

$$\lim_{T\to\infty}p_t(T,v)=\lim_{T\to\infty}p_t(T-1,v) \tag{3.86}$$

But also notice that given the stationary structure of the equilibrium

$$p_t(T-1,v)=p_{t+1}(T,v) \tag{3.87}$$

for all t and T. Hence we can say

$$\lim_{T\to\infty}p_t(T,v)=\lim_{T\to\infty}p_{t+1}(T,v) \tag{3.88}$$

This given the structure of the T-period equilibrium implies that

$$\lim_{T\to\infty}y_t(T)=\lim_{T\to\infty}y_{t+1}(T) \tag{3.89}$$

We can apply the same analysis to the qualities which would give

$$\lim_{T\to\infty}x_t(T)=\lim_{T\to\infty}x_{t+1}(T) \tag{3.90}$$

But then when we apply these limits to the identifying equations of the T-period equilibrium we get

$$\left(2\lim_{T\to\infty} x_t(T) - \lim_{T\to\infty} y_t^2(T)\right)(2 - 3\lim_{T\to\infty} y_t(T)) + \delta \lim_{T\to\infty} y_t^3(T)(2\lim_{T\to\infty} x_{t+1}(T) - \lim_{T\to\infty} y_{t+1}^2(T)) = 0$$

But since

$$\left(2\lim_{T\to\infty} x_t(T) - \lim_{T\to\infty} y_t^2(T)\right) = (2\lim_{T\to\infty} x_{t+1}(T) - \lim_{T\to\infty} y_{t+1}^2(T)) \tag{3.91}$$

this reduces the equation to

$$\delta \lim_{T\to\infty} y_t^3(T) - 3\lim_{T\to\infty} y_t(T)) + 2 = 0 \tag{3.92}$$

But we know that the solution to this equation is given by

$$\lim_{T\to\infty} y_t(T) = y(\delta) \tag{3.93}$$

which proves the convergence. ∎

Having shown that the equilibrium we proposed here is the limiting equilibrium of the finite period game, we can discuss some of its properties. First let us look at the equilibrium path behavior of the seller and the buyers.

Corollary 3.1. *The following represent the behavior on the equilibrium path:*

1. $p_t^*(\delta) = x(\delta)\, y(\delta)^{2t}$
2. $q_t^*(\delta) = y(\delta)^{t+1}$

3. $v_t^*(\delta) = y(\delta)^t$

4. $\Pi_0^*(\delta) = \frac{(1-y(\delta))\left(x(\delta)-\frac{1}{2}y(\delta)^2\right)}{1-\delta y(\delta)^3}$ *where* $x(\delta)$ *and* $y(\delta)$ *are as defined in Proposition 3.3.*

Having found the equilibrium path behavior, we can compare it to what happens in a fixed quality model. First of all notice that if the quality is fixed at zero the buyers become homogenous, hence we can restrict attention to strictly positive levels of quality. Notice that for all non-degenerate levels of fixed quality, there will be a strictly positive marginal cost. But this means the seller will not cover the whole market but rather sell up to the buyer with the total valuation equal to the marginal cost.

Sobel and Takashi (1983) find a stationary equilibrium of the same model but with the quality fixed at unity and with zero marginal cost. The equilibrium they find is also the limit for the finite period game. Suppose we fix the quality at $\overline{q}$, this means the marginal cost is $\frac{\overline{q}^2}{2}$ and hence the last buyer to be served has a valuation $\frac{\overline{q}}{2}$. When we apply their finding to this setting we get the following equilibrium:

$$\widehat{p}_t^*(\widehat{v}_t) = \gamma(\delta)\,\widehat{v}_t \tag{3.94}$$

$$\widehat{v}_{t+1}^*(p_t) = \lambda(\delta)\,\widehat{p}_t \tag{3.95}$$

where $\widehat{v}_t$ and $\widehat{p}_t$ represents the valuation and prices net of marginal cost $\frac{\overline{q}^2}{2}$, i.e. the buyer with redefined valuation $\widehat{v}_t$ corresponds to buyer with actual type $v_t = \frac{1}{\overline{q}}\left(\widehat{v}_t + \frac{\overline{q}^2}{2}\right)$. Given these interpretations then we can simply take Sobel and Takashi's solution for $\lambda(\delta)$ and

$\gamma(\delta)$, which are

$$\lambda(\delta) = \frac{1}{\sqrt{1-\delta}} \quad (3.96)$$

$$\gamma(\delta) = \frac{\sqrt{1-\delta}-(1-\delta)}{\delta} \quad (3.97)$$

Using these expressions we can see that the equilibrium path is described by the following:

$$\widehat{v}_t^*(\delta,\overline{q}) = \gamma(\delta)^t \lambda(\delta)^t \left(\overline{q}-\frac{\overline{q}^2}{2}\right) \quad (3.98)$$

$$\widehat{p}_t^*(\delta,\overline{q}) = \gamma(\delta)^{t+1} \lambda(\delta)^t \left(\overline{q}-\frac{\overline{q}^2}{2}\right) \quad (3.99)$$

$$\widehat{\Pi}_0^*(\delta,\overline{q}) = \frac{\gamma(\delta)\left(1-\gamma(\delta)\lambda(\delta)\right)\left(\overline{q}-\frac{\overline{q}^2}{2}\right)^2}{\left(1-\delta\gamma(\delta)^2\lambda(\delta)^2\right)\overline{q}} \quad (3.100)$$

where $\widehat{\Pi}_0^*(\delta,\overline{q})$ represents equilibrium profits given the fixed quality. Notice that $\widehat{p}_t^*(\delta,\overline{q})$ is the mark-up in this setting. The rate of change of the mark-up is then given by $\ln(\gamma(\delta)\lambda(\delta))$. We can compare this to the rate of change of the mark-up in the variable quality model, which by Corollary 3.1, is given by $\ln\left(y(\delta)^2\right)$. An algebraic comparison reveals

$$y(\delta)^2 < \gamma(\delta)\lambda(\delta) < 1$$

which means that the profit margin declines at a smaller rate when the quality is fixed. Notice that the rate of change of prices is the same as that of the profit margin. Hence this also means that the variable quality model implies a higher rate of decrease in prices.

Going back to the profit expression $\widehat{\Pi}_0^*(\delta,\overline{q})$, the optimal fixed quality for the seller then can be found by maximizing $\frac{\left(\overline{q}-\frac{\overline{q}^2}{2}\right)^2}{\overline{q}}$. The maximizer of this function is $\overline{q} = \frac{2}{3}$, which is the optimal quality for the static monopolist. Now we can compare the mark-up, price and quality at which the market opens using this optimal fixed quality. Using the expressions in Corollary 3.1 and the definitions of $x(\delta)$ and $y(\delta)$ in Proposition 3.3 careful algebraic comparison shows that

$$\widehat{p_0^*}\left(\delta, \frac{2}{3}\right) < p_0^*(\delta) - \frac{1}{2} q_0^*(\delta)^2$$

for all δ, which means that the market opens with a smaller mark-up when the quality is fixed at the optimal level than when it is variable. A similar comparison would also show that

$$\begin{aligned} \widehat{p_0^*}\left(\delta, \frac{2}{3}\right) + \frac{2}{9} &< p_0^*(\delta) \\ \frac{2}{3} &< q_0^*(\delta) \end{aligned}$$

which means that the higher mark-up at period 0 is achieved through a higher price and quality in the case of variable quality.

We can also look at the difference between the overall profits. It could be algebraically shown that

$$\widehat{\Pi_0^*}\left(\delta, \frac{2}{3}\right) < \Pi_0^*(\delta)$$

for all δ. This means the seller is actually better off not committing to a fixed quality. The reason for this is that, as we explained in the two-period model, the seller uses the different time periods as a way to offer a menu of contracts with different price and quality pairs and hence extracts more surplus.

3.4.1 Properties of the equilibrium

In this section we will look at some comparative static and limiting properties of the stationary equilibrium that we described in Proposition 3.3.

Comparative static properties

We will start by looking at how $y(\delta)$ and $x(\delta)$ vary in δ, which through Corollary 3.1 will help understand how the equilibrium behavior changes.

Proposition 3.5. *$y(\delta)$ and $x(\delta)$ increase in δ.*

Proof. Recall that $y(\delta)$ is defined by the polynomial in Proposition 3. Applying implicit function theorem to that polynomial we get

$$\frac{dy(\delta)}{d\delta} = \frac{y(\delta)^3}{3(1-\delta y(\delta)^2)} \geqslant 0 \tag{3.101}$$

Now let us look at $x(\delta)$, which is defined in the same proposition.

$$\begin{aligned} \frac{dx(\delta)}{d\delta} &= \frac{\partial x(\delta)}{\partial \delta} + \frac{\partial x(\delta)}{\partial y(\delta)}\frac{dy(\delta)}{d\delta} \\ &= y(\delta)^3 \frac{-1+y(\delta)}{\left(-1+\delta r y(\delta)^2\right)^2} + \left(y(\delta)\frac{2-3\delta y(\delta)+\delta^2 y(\delta)^3}{\left(-1+\delta y(\delta)^2\right)^2}\right)\frac{y(\delta)^3}{3\left(1-\delta y(\delta)^2\right)} \\ &= \frac{y(\delta)^3\left(\delta^2 y(\delta)^4 - 3\delta y(\delta)^3 + 5y(\delta) - 3\right)}{3\left(1-\delta y(\delta)^2\right)^3} \geqslant 0 \end{aligned} \tag{3.102}$$

because it can be seen that for $y \leqslant 1$ that satisfy $\delta y^3 - 3y + 2 = 0$, $r^2y^4 - 3ry^3 + 5y - 3 \geqslant 0$. ∎

Proposition 3.5 together with Corollary 3.1 says that each period's price and quality increases in δ and also that the proportional variation in the price and quality across periods decreases. In other words price and quality decline at a smaller rate as δ increases. This also implies that the buyers are partitioned more finely over time for higher δ.

If we also look at the profit margin on the equilibrium path for time t, using Corollary 3.1 it is given by

$$p_t^* - \frac{1}{2}q_t^{*2} = \left(x(\delta) - \frac{1}{2}y(\delta)^2\right) y(\delta)^{2t} \tag{3.103}$$

First of all at the equilibrium levels of $y(\delta)$ and $x(\delta)$, the term in parenthesis can be shown to be positive. Notice also that as δ increases, due to the increase in $y(\delta)$, the profit margin declines at a smaller rate. The intuition behind this result is the following: As the players become more patient, they are willing to wait longer for more desirable offers. For them not to wait the rate of decrease in price has to become much lower than the rate of decrease in quality. But this would imply that the profit margin declines at a smaller rate.

Looking at how the profit margin at each t changes with δ would not yield much information, because the group of buyers that purchase at each t also change with δ. But we can look at the profit margin the seller opens the market with.

Proposition 3.6. *The equilibrium profit margin at $t = 0$ decreases in δ.*

Proof. The equilibrium profit margin at $t = 0$ is given by

$$x(\delta) - \frac{1}{2}y(\delta)^2 = \frac{y(\delta)^2\left(\delta y(\delta)^2 - 2\delta y(\delta) + 1\right)}{2\left(1 - \delta y(\delta)^2\right)} \tag{3.104}$$

Taking the derivative with respect to δ yields

$$\begin{aligned}
\frac{d\left(x(\delta)-\frac{1}{2}y^2\right)}{d\delta} &= \frac{\partial\left(x(\delta)-\frac{1}{2}y(\delta)^2\right)}{\partial\delta}+\frac{\partial\left(x(\delta)-\frac{1}{2}y(\delta)^2\right)}{\partial y(\delta)}\frac{dy(\delta)}{d\delta} \\
&= y(\delta)^3\frac{-1+y(\delta)}{\left(-1+\delta y(\delta)^2\right)^2}+ \\
&\left(-y(\delta)\frac{-1+3\delta y(\delta)-\delta^2 y(\delta)^3+\delta^2 y(\delta)^4-2\delta y(\delta)^2}{\left(-1+\delta y(\delta)^2\right)^2}\right)\frac{y(\delta)^3}{3\left(1-\delta y(\delta)^2\right)} \\
&= \frac{y(\delta)^3\left(-\delta^2 y(\delta)^5+\delta^2 y(\delta)^4-\delta y(\delta)^3+4y(\delta)-3\right)}{3\left(1-\delta y(\delta)^2\right)^3}\leqslant 0
\end{aligned} \tag{3.105}$$

because it can be seen that for $y(\delta)\leqslant 1$ that satisfy $\delta y(\delta)^3-3y(\delta)+2=0$, $-r^2y(\delta)^5+r^2y(\delta)^4-ry(\delta)^3+4y(\delta)-3\leqslant 0$. ∎

Proposition 3.6 says that for high δ the seller opens the market with a higher price and quality but a lower margin. Recall that we argued above that the profit margin declines at a slower rate for high δ. But due to the transversality condition we know that at the limit, as $T\to\infty$, the profit margin vanishes for all δ. This implies that for high δ, the seller should start with a lower margin.

Now let us look at the infinite-period profits as defined in Corollary 3.1.

Proposition 3.7. $\Pi_0^*(\delta)$ *decreases in* δ.

Proof. Substituting for $x(\delta)$ in the expression given in Corollary 3.1 results in

$$\Pi_0^*(\delta)=\frac{1}{2}(1-y(\delta))\,y(\delta)^2\frac{1-2\delta y(\delta)+\delta y(\delta)^2}{\left(1-\delta y(\delta)^2\right)\left(1-\delta y(\delta)^3\right)} \tag{3.106}$$

Using equation (3.101) we get

$$\frac{d\Pi_0^*(\delta)}{d\delta} = -\frac{1}{6}y^3 \frac{\begin{pmatrix} 6 - 14y + 6y^2 - 2\delta^2 y^4 + 3y^3 - 6\delta^3 y^8 \\ +2\delta^3 y^7 - y^7\delta^2 - 9y^5\delta + 3y^9\delta^3 + 6y^4\delta + 6y^6\delta^2 \end{pmatrix}}{(1-\delta y^2)^3 (1-\delta y^3)^2} \tag{3.107}$$

It can be shown algebraically that for $y(\delta) \leqslant 1$ that satisfy $\delta y(\delta)^3 - 3y(\delta) + 2 = 0$, the numerator is positive, which makes the whole term negative. ▮

The last proposition implies that the seller makes less profits as the players become more patient.

Limiting properties

Having looked how the equilibrium changes with δ, now we will investigate where it converges as the players become infinitely patient. Again let us first start with the definitive elements of the equilibrium, namely $y(\delta)$ and $x(\delta)$.

Proposition 3.8. $\lim_{\delta \to 1} y(\delta) = 1$ *and* $\lim_{\delta \to 1} x(\delta) = \frac{1}{2}$.

Proof. We know that $y(\delta)$ is the solution to the polynomial given in Proposition 3.3, which implies that it is continuous. Given that it is continuous, we can simply take the limits of the left and right hand sides of the polynomial in Proposition 3.3, which gives

$$\begin{aligned} \left(\lim_{\delta \to 1} \delta\right) \left(\lim_{\delta \to 1} y(\delta)\right)^3 - 3\left(\lim_{\delta \to 1} y(\delta)\right) + 2 &= 0 \\ \left(\lim_{\delta \to 1} y(\delta)\right)^3 - 3\left(\lim_{\delta \to 1} y(\delta)\right) + 2 &= 0 \end{aligned} \tag{3.108}$$

This is a polynomial in $\lim_{\delta \to 1} y(\delta)$. The unique solution that satisfies the second condition in Proposition 3.3 is $\lim_{\delta \to 1} y(\delta) = 1$.

Now let us look at $x(\delta)$. Notice that

$$x(1) = \frac{0}{0} \tag{3.109}$$

so we can use L'Hospital Rule to take the limit, i.e.

$$\begin{aligned}
\lim_{\delta \to 1} x(\delta) &= \lim_{\delta \to 1} \frac{\frac{d}{d\delta}\left(y(\delta)^2 - \delta y(\delta)^3\right)}{\frac{d}{d\delta}\left(1 - \delta y(\delta)^2\right)} \\
&= \lim_{\delta \to 1} \frac{-y(\delta)^3 + \left(2y(\delta) - 3\delta y(\delta)^2\right)\frac{dy(\delta)}{d\delta}}{-y(\delta)^2 - 2\delta y(\delta)\frac{dy(\delta)}{d\delta}} \\
&= \lim_{\delta \to 1} \frac{-y(\delta)^3 + \left(2y(\delta) - 3\delta y(\delta)^2\right)\frac{dy(\delta)}{d\delta}}{-y(\delta)^2 - 2\delta y(\delta)\frac{dy(\delta)}{d\delta}} \\
&= \lim_{\delta \to 1} \frac{\frac{-y(\delta)^3}{\frac{dy(\delta)}{d\delta}} + \left(2y(\delta) - 3\delta y(\delta)^2\right)}{\frac{-y(\delta)^2}{\frac{dy(\delta)}{d\delta}} - 2\delta y(\delta)} \\
&= \frac{\frac{-1}{\lim_{\delta \to 1}\frac{dy(\delta)}{d\delta}} - 1}{\lim_{\delta \to 1}\frac{-1}{\frac{dy(\delta)}{d\delta}} - 2} \\
&= \frac{1}{2}
\end{aligned} \tag{3.110}$$

because

$$\lim_{\delta \to 1} \frac{dy(\delta)}{d\delta} = \lim_{\delta \to 1}\left(\frac{y(\delta)^3}{3(1 - \delta y(\delta)^2)}\right) = \infty \tag{3.111}$$

∎

The fact that $\lim_{\delta \to 1} y(\delta) = 1$ implies two things: (1) the market opens with quality 1 in the limit and (2) the buyers are perfectly partitioned as $\delta \to 1$. This means the menu of contracts offered perfectly separate each buyer at the limit. However the fact that $\lim_{\delta\to 1} x(\delta) = \frac{1}{2}$ means that the seller is opening the market with price $\frac{1}{2}$ and hence zero profit margin. Therefore the seller gives up all the surplus as the following corollary suggests.

Corollary 3.2. *On the equilibrium path, the profit margin each period converges to zero as* $\delta \to 1$.

Proof. Equilibrium path profit margin in period t is given by equation (3.103). Taking the limit of both sides gives

$$\begin{aligned}\lim_{\delta\to 1}\left(p_t^*(\delta) - \frac{1}{2}q_t^*(\delta)^2\right) &= \left(\lim_{\delta\to 1} x(\delta) - \frac{1}{2}\left(\lim_{\delta\to 1} y(\delta)\right)^2\right)\left(\lim_{\delta\to 1} y(\delta)\right)^{2t} \quad (3.112)\\ &= \left(\frac{1}{2} - \frac{1}{2}\right)\\ &= 0\end{aligned}$$

which proves the proposition. ∎

We know that the profit margin each period goes to zero. Now let us look at the volume of sales each period.

Proposition 3.9. *The equilibrium path volume of sales each period goes to zero as* $\delta \to 1$.

Proof. Recall from Corollary 3.1 that $v_t^*(\delta) = y(\delta)^t$, which means the volume of sales in period t is

$$\begin{aligned}v_t^*(\delta) - v_{t+1}^*(\delta) &= y(\delta)^t - y(\delta)^{t+1} \quad (3.113)\\ &= y(\delta)^t(1 - y(\delta))\end{aligned}$$

Taking the limit of both sides yields

$$\lim_{\delta\to 1}\left(v_t^*(\delta) - v_{t+1}^*(\delta)\right) = 0 \tag{3.114}$$

which proves the proposition. ∎

Not only does the profit margin disappear but also the measure of sales become infinitesimally small as the players become infinitely patient. However, this does not automatically mean that the infinite period profits vanish. There is the possibility that almost zero profits each period might add up to a non-zero infinite-period value. To see that we need to look at what $\Pi_0^*(\delta)$ converges to as $\delta \to 1$. The next proposition shows that $\Pi_0^*(\delta)$ indeed converges to zero.

Proposition 3.10. $\lim_{\delta\to 1}\Pi_0^*(\delta) = 0$

Proof. Using equation (3.106) notice that $\Pi_0^*(1) = \frac{0}{0}$ so we can apply L'Hospital's rule to find the limit. This implies that

$$\begin{aligned}
\lim_{\delta\to 1}\Pi_0^*(\delta) &= \lim_{\delta\to 1}\frac{\frac{d}{d\delta}\left(\left(1-y(\delta)\right)y(\delta)^2\left(1-2\delta y(\delta)+\delta y(\delta)^2\right)\right)}{\frac{d}{d\delta}\left(2\left(1-\delta y(\delta)^2\right)\left(1-\delta y(\delta)^3\right)\right)} \\
&= \lim_{\delta\to 1}\frac{\begin{array}{c}-\left(-1+y(\delta)\right)y(\delta)^3\left(-2+y(\delta)\right)\\ -y(\delta)\left(3y(\delta)-12\delta y(\delta)^2+5y(\delta)^3\delta-2+6y(\delta)\delta\right)\frac{dy(\delta)}{d\delta}\end{array}}{\begin{array}{c}2y(\delta)^2\left(-1+2y(\delta)^3\delta-y(\delta)\right)\\ +2y(\delta)\delta\left(-2+5y(\delta)^3\delta-3y(\delta)\right)\frac{dy(\delta)}{d\delta}\end{array}} \\
&= \lim_{\delta\to 1}\frac{1}{2}y(\delta)\frac{6-11y(\delta)-3\delta y(\delta)^3+6y(\delta)^2+2y(\delta)^4\delta}{3-\delta y(\delta)^2-6\delta y(\delta)^3+\delta^2 y(\delta)^5+3y(\delta)}
\end{aligned} \tag{3.115}$$

where the third line follows from substitution of $\frac{dy(\delta)}{d\delta}$ from equation (3.101). Notice that the last expression also becomes $\frac{0}{0}$ at $\delta = 1$, so we can apply L'Hospital one more time to get

$$\lim_{\delta \to 1} \Pi_0^* (\delta) = \lim_{\delta \to 1} \frac{\begin{array}{c} -3y(\delta)^4 + 2y(\delta)^5 \\ + \left(6 - 22y(\delta) - 12\delta y(\delta)^3 + 18y(\delta)^2 + 10y(\delta)^4 \delta\right) \frac{dy(\delta)}{d\delta} \end{array}}{\begin{array}{c} -2y(\delta)^2 - 12y(\delta)^3 + 4y(\delta)^5 \delta \\ + \left(-4\delta y(\delta) - 36\delta y(\delta)^2 + 10\delta^2 y(\delta)^4 + 6\right) \frac{dy(\delta)}{d\delta} \end{array}} \tag{3.116}$$

$$= \lim_{\delta \to 1} -\frac{1}{2} y(\delta) \frac{-31y(\delta) - 3\delta y(\delta)^3 + 24y(\delta)^2 + 4y(\delta)^4 \delta + 6}{3 - \delta y(\delta)^2 + 15y(\delta) - 6\delta y(\delta)^3 + \delta^2 y(\delta)^5}$$

$$= 0$$

which proves the proposition. ∎

The last proposition says that as the players become infinitely patient, the seller's overall profits vanish. We can directly interpret the discount rate going to one as the offers being made very frequently, if we define the discount rate as $\delta = e^{-r\Delta}$ where r is the interest rate and Δ is the real time between two offers. This means as the offers take place very quickly, i.e. as $\Delta \to 0$ the seller's profit converges to zero. This part of the Coasian conjecture is verified. However, Coase conjecture has another implication which is that the market becomes more and more efficient and the buyers get all the surplus as the offers are made very frequently. To answer these questions we need to be careful. First of all the concept of efficiency is different in this model compared to the fixed quality model. In the fixed quality model efficiency refers to serving the buyers with valuations higher than the marginal cost. However, here efficient allocation actually consists of supplying the optimal quality product to each buyer. In a static world if a social planner knew the types of each buyer then he would choose personalized qualities for each v, $q(v)$, to maximize the following social surplus

$$\max_{q(v)} \int_0^1 vq(v) - \frac{1}{2}q(v)^2 \tag{3.117}$$

Piecewise maximization of the social planner's surplus yields

$$q(v) = v \tag{3.118}$$

as the socially optimal quality for each buyer. We know that as the time between offers gets close to zero, offers are made almost in real time. We showed that, the price and quality given market penetration v approaches to $\frac{1}{2}v$ and v respectively because $y(\delta)$ converges to 1 and $x(\delta)$ converges to $\frac{1}{2}$ as $\delta \to 1$ and hence also as $\Delta \to 0$. But if the offers are almost made in real time that means the buyer with actual type v is almost getting the offer with price $\frac{1}{2}v$ and quality v. But this is the socially optimal quality for type v. This proves convergence to efficiency. This argument also reveal how the surplus is shared. Notice that the buyer with type v almost gets quality v at price $\frac{1}{2}v$, which yields a zero surplus for the seller for each buyer. This means that just like in the fixed quality case the buyers are getting almost all the surplus.

3.5 Conclusion

We showed that Coase conjecture still applies when the seller can vary quality over time along with the price. We verified this result for one particular equilibrium, which is the limiting equilibrium of the finite-period game. In this equilibrium, we showed that as the time between offers gets close to zero, the seller's profits vanish and the buyers enjoy all the surplus with the efficient allocation. The efficient allocation corresponds to a menu of contracts that offer the individual surplus maximizing quality to each buyer.

We also compared the seller's profits between the variable quality and optimal fixed quality scenarios. We found that the seller enjoys higher profits when he can vary the quality and hence offer a menu of contracts over time. This result is not surprising, because quality variation allows the seller to separate out buyers with different willingness to pay and hence extract more surplus. However, when the offers are made very frequently, the seller's profits collapse in both scenarios.

The deriving force behind the collapse of the profits is the "skimming property" that we impose on the equilibrium relying on the fact that it holds for the finite game. The skimming property, i.e. the willingness of the seller to move down the demand curve, is so strong that it generates the same effects as in a fixed quality model. The only difference here is that it happens not solely through declining prices but also through declining quality.

One can claim that there are non-stationary equilibria of this game that might violate the Coase conjecture. This is also a limitation of Coase conjecture when applied to a fixed quality game. The more interesting violation of Coase conjecture could be through a stationary equilibrium that violates the skimming property. We cannot rule out or verify the existence of such equilibria. For further research, it would be interesting to analyze the existence and characterization of such equilibria.

Bibliography

ADAMS, W. J. AND J. L YELLEN (1976): "Commodity bundling and the burden of monopoly," *Quarterly Journal of Economics*, 90, 475-98.

ANSARI, A.,S. ESSAGAIER, AND R. KOHLI (2000): "Internet Recommendation Systems", *Journal of Marketing Research,* 37, 363-375.

AUSUBEL, L., AND R. DENECKERE (1989): "Reputation in Bargaining and Durable Goods Monopoly", *Econometrica*, 57, 511-531.

AVERY, C., P. RESNICK, AND R. ZECKHAUSER (1999): "The Market for Evaluations", *American Economic Review*, 89, 564-584.

AVERY, C., AND R. ZECKHAUSER (1997), "Recommender systems for evaluating computer messages", *Communications of the ACM, 40, 88-89.*

BREESE, J.,S., D. HECKERMAN,AND C. KADIE (1998): "Empirical Analysis of Predictive Algorithms for Collaborative Filtering", *Technical Report MSR-TR-98-12*, Microsoft Research

BRYNJOLFSSON E. AND M. D. SMITH (2001): "The Great Equalizer? Consumer Choice Behavior at Internet Shopbots", *Sloan Working Paper 4208-01.*

BULOW, J. (1982): "Durable Good Monopolists," *Journal of Political Economy*, 90, 314-322.

CHEVALIER, J., AND D. MAYZLIN (2003): "The Effect of Word of Mouth on Sales: Online Book Reviews", *Yale SOM Working Papers,* ES-28 & MK-15.

CRAWFORD V. AND J. SOBEL (1982): "Strategic Information Transmission", *Econometrica*, 50, 1431-1451.

ECONOMIDES, N. (1996): "The Economics of Networks", *International Journal of Industrial Organization*, 16, 271-284.

ECONOMIDES, N. AND C. HIMMELBERG (1995): "Critical Mass and Network Size with Application to the US FAX Market", *Stern School of Business Discussion Papers,* No. EC-95-11.

EPPEN, G. D., W. A. HANSON AND R. K. MARTIN (1991): "Bundling-New Products, New Markets, Low Risk.", *Sloan Management Review,* 32, 7-14.

FARRELL J. AND R. GIBBONS (1989): "Cheap Talk with Two Audiences", *The American Economic Review*, 79, 1214-1223.

FUDENBERG, D., D. LEVINE,AND J. TIROLE (1985): "Infinite Horizon Models of Bargaining with One-Sided Incomplete Information", in *Game Theoretic Models of Bargaining,* ed. A. Roth. Cambridge University Press.

GUL, F., H. SONNENSCHEIN AND R. WILSON (1986): "Foundations of Dynamic Monopoly and the Coase Conjecture", *Journal of Economic Theory,* 39, 155-190.

IEEE INTERNET COMPUTING: INDUSTRY REPORT: "Amazon.com Recommendations: Item-to-Item Collaborative Filtering", *http://dsonline.computer.org/0301/d/w1lind.htm.*

KARP, L. (1996): "Depreciation Erodes the Coase Conjecture", *European Economic Review*, 40, 473-490.

MCAFEE, P., AND T. WISEMAN (2004): "Capacity Choice Counters the Coase Conjecture", *mimeo.*

MILD, A. AND M. NATTER (2001): "A Critical View on Recommendation Systems", *Working Papers SFB "Adaptive Information Systems and Modelling in Economics and Management Science"*, Nr. 82

MILGROM P. AND J. ROBERTS (1986): "Relying on Information of Interested Parties", *The RAND Journal of Economics*, 17, 18-32.

RESNICK, P. AND H. VARIAN (1997): "Recommender Systems", *Communications of the ACM, 40, 56-58.*

SCHMALENSEE, R. L. (1984): "Gaussian demand and commodity bundling.", *Journal of Business*, 57, 211-230.

SHAPIRO, C. AND R. VARIAN (1999): *Information Rules, A Strategic Guide to the Network Economy.* Harvard: Harvard Business School Press.

SHIN H. S. (2003): "Disclosures and Asset Returns", *Econometrica*, 71, 105-133.

SOBEL, J., AND I. TAKAHASHI (1983), "A Multi Stage Model of Bargaining", *Review of Economic Studies, 50, 411-426.*

STOKEY, N. (1981): "Rational Expectations and Durable Goods Pricing", *Bell Journal of Economics,* 12, 112-128.

VULKAN, N. (2003), *The Economics of E-Commerce, A Strategic Guide to Understanding and Designing the Online Marketplace.* Princeton: Princeton University Press.