OPTIMAL PRICING WITH RECOMMENDER SYSTEMS

By

DIRK BERGEMANN AND DERAN OZMEN

March 2006

COWLES FOUNDATION DISCUSSION PAPER NO. 1563

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Optimal Pricing with Recommender Systems*

Dirk Bergemann
Yale University
New Haven, CT 06511
dirk.bergemann@yale.edu

Deran Ozmen
Yale University
New Haven, CT 06511
deranozmen@gmail.com

March 2006
EXTENDED ABSTRACT

Abstract

We study optimal pricing in the presence of recommender systems. A recommender system affects the market in two ways: (i) it creates value by reducing product uncertainty for the customers and hence (ii) its recommendations can be offered as add-ons which generate informational externalities. The quality of the recommendation add-on is endogenously determined by sales. We investigate the impact of these factors on the optimal pricing by a seller with a recommender system against a competitive fringe without such a system.

If the recommender system is sufficiently effective in reducing uncertainty, then the seller prices otherwise symmetric products differently to have some products experienced more aggressively. Moreover, the seller segments the market so that customers with more inflexible tastes pay higher prices to get better recommendations.

KEYWORDS: Recommender System, Collaborative Filtering, Add-Ons, Pricing, Information Externality

JEL CLASSFICATION: D42, D83, D85.

1 Introduction

The increased use of the internet for commercial transactions leads to a large accumulation of data about customers and products on the internet. In particular, the 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

*The first author acknowledges financial support through NSF #CNS 0428422.and #SES-0518929. The extended abstract was prepared for ACM-EC 2006.

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. In this paper we focus on the role of prices by assuming that each buyer has unit demand in each period. 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.

The recommendation can be considered as an add-on: it is an additional service a customer receives on top of the purchase she makes. A similar interpretation along 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. We find that when the recommender system reduces uncertainty only by a small amount, then 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 book 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

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

able to use the information provided by the large group. Those with the low willingness to pay choose to be in the larger 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 book 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 to 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.

The road map is as follows: We discuss the related literature in Section 2. The model is described in Section 3. In Section 4 we analyze the optimal pricing policy and the resulting equilibrium allocation. Section 5 concludes by discussing possible extensions for future research.

2 Related Literature

The possibility of increasing sales through recommender systems 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.

Varian and Resnick (1997) give a brief description of recommender systems and the issues they raise. They 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".

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

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

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.

In computer science, the recommender systems we discuss here are formally known as "collaborative filtering systems". The "IEEE Internet Computing: Industry Report", describes that Amazon.com uses a modified collaborative filtering method referred to as the "item-to-item based collaborative filtering". This method computes a 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 classic linear regression models. 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.

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. A buyer's product choice In period 1, each buyer decides between a new product and a product which she did not purchase in period 0:

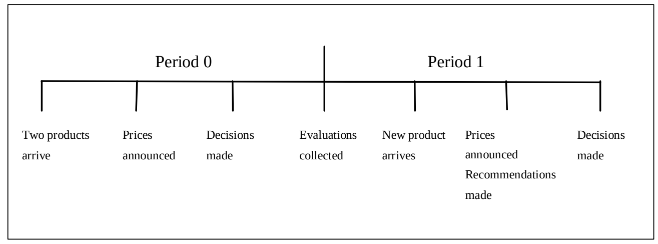

Figure 1: The timeline

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. 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

$$u(\theta, x) = v - (\theta - x)^2 . \quad (1)$$

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.

The quadratic utility model is chosen for computational ease and any utility function which decreases in the distance between the ideal product type and the current product type would generate the same qualitative results.

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 uncertainty symmetric around $\frac{1}{2}$:

$$\alpha_l = \frac{1}{2} - \varepsilon, \qquad \alpha_h = \frac{1}{2} + \varepsilon, \tag{2}$$

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". The symmetry permits us to represent the initial information by a one-dimensional parameter, but the symmetry by itself is not essential for either the analysis or the results to come.

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 what the exact value of α_m will be in period 1, but they attach $\frac{1}{2}$ probability to α_m being α_h and α_l. The products l and h continue to be available in period 1.

The marginal cost of each product for all sellers is c. We assume that the price for each product in the competitive fringe equals c. Each buyer buys at most one product each period. Moreover, a buyer wishes to buy a different product each period. We assume that per period outside utility for each buyer is smaller than $v - c - 4$, so that absent a better offer from the seller with the recommender system, each buyer is willing to buy one product in each period from the fringe.

Learning 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 $y_i(x_i) \in \{-1, 0, 1\}$ on the type of each product $i \in \{l, h\}$ between periods 0 and 1, where

$$\begin{aligned} \Pr(y_i(x_i) = 0 \mid x_i) &= 1 - \mu_i, \\ \Pr(y_i(x_i) \in \{-1, 1\} \mid x_i) &= \mu_i. \end{aligned}$$

We can interpret a signal of 0 as containing no information, or simply the failure to receive an informative 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 $\gamma \in \left[0, \frac{1}{2}\right]$. We can interpret γ as the informativeness of the signal. The event tree in Figure 2 summarizes the signal structure where $x_i' \neq x_i$.

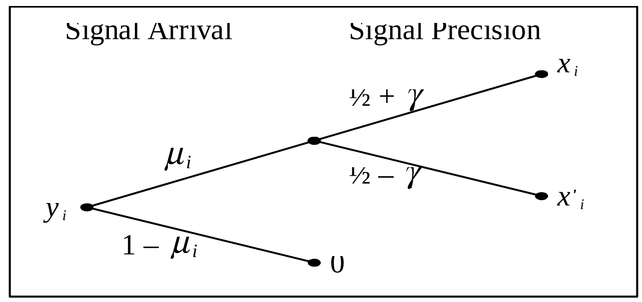

Figure 2: The Signal Structure

Given the probabilistic structure, we view the recommender system as a 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

$$\alpha_i(y_i) \equiv \Pr(x_i = 1 \mid y_i).$$

We assume that only buyers which bought from seller M benefit from the information of seller $\dot{M}$. This might seem to be important restriction relative to the current practice of some recommender systems who provide information even to new customers. However, we observe that all we need for the pricing model here is that past customers receive statistically more valuable information than new customer. In many recommender systems this occurs through personalized recommendations.

Pricing In period 0, seller M announces prices for each product, i.e. $\mathbf{p} = (p_l, p_h) \in \mathbb{R}^2$. 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, $(p'_l, p'_m, p'_h) \in \mathbb{R}^3$ and reveals the recommendations to the buyers who have purchased from him in period 0. The recommendation of a buyer who purchased $i \in \{l, h\}$ is a report about the product she did not purchase yet, i.e. an estimate about the value of $j \equiv \{l, h\} / i$. Given the recommendation and the prices, each buyer then decides whether to purchase the new product or the product she did not purchase in period 0. 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.

Interpretation 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 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

$$\rho = \frac{\gamma}{\varepsilon}, \tag{3}$$

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.

Figure 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.

4 Equilibrium

In this section we investigate the Perfect Bayesian Equilibria of the game between seller M and the buyers. The seller M announces a price for each product in period 0 and each buyer optimally chooses a seller and a product given the prices. It is entirely straightforward to find the equilibrium 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 F are effectively selling identical products. The competition is fierce and the equilibrium price will have to equal the marginal cost c. When we introduce some uncertainty and informativeness into the setting, the distribution of buyers in period 0 affects the information gathered and hence the utilities in period 1. As seller M has the sole control over the distribution of buyers in period 0, he controls how much information is gathered for period 1. In period 1, seller M then reveals the information he has gathered 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 to collect information. We will start analyzing M's problem with the subgame in period 1.

Lemma 1 (PERIOD 1 SUBGAME)
The minimum price in the market in period 1 in any perfect Bayesian equilibrium equals marginal cost for each product.

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 $\mathbf{p} = (p_h, p_l)$. Given the difference between the two prices the distribution of buyers could be different. The following two properties of the distribution of buyers over products will be central for the determination of the equilibrium.

Definition 1 (BALANCE)
A distribution (μ_h, μ_l) of buyers is balanced if $\mu_h = \mu_l$ and unbalanced if $\mu_i > \mu_j$ for some $i \in \{l, h\}$. The degree of imbalance is $\frac{\mu_i}{\mu_j}$.

We would like to see whether the seller creates endogenous differentiation between the two products through an unbalanced distribution and if so, which buyers 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 to know the composition of these groups. This suggests the following definition.

Definition 2 (SORTING)
A distribution (μ_l, μ_h) of buyers 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. We refer to*
$$\frac{\min\{|S^-|, |S^+|\}}{\max\{|S^-|, |S^+|\}}$$
as the degree of shuffling;
3. *"perfectly shuffled" if $|S^-| = |S^+|$;*
4. *"with a gap" if the distribution satisfies either of above criteria except that a segment around zero separates either the buyers of j into two segments or separates the buyers of i and j.*

Figure 3 and 4 illustrate Definition 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.

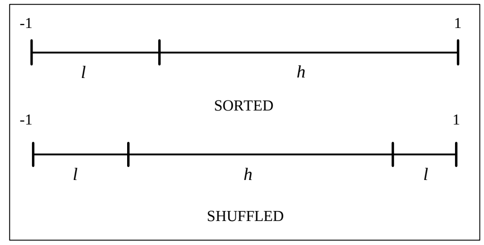

Figure 3: Sorting and Shuffling for $\mu_h > \mu_l$

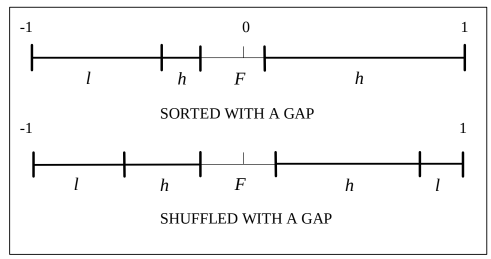

Figure 4: Sorting and Shuffling with a Gap

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

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

First, as we discussed above,in period 1 all the products are sold at marginal cost so the buyers make their choice to maximize their expected utility with respect to the remaining choice set $\{j, m\}$, given the recommendations. Second, notice that equation (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

$$\begin{aligned} \overline{\alpha}_i\left(\mu_j\right) &\equiv \mathbb{E}\left(\max\left\{\alpha_m, \alpha_j\left(y_j\right)\right\} \mid \mu_j\right), \\ \underline{\alpha}_i\left(\mu_j\right) &\equiv \mathbb{E}\left(\min\left\{\alpha_m, \alpha_j\left(y_j\right)\right\} \mid \mu_j\right). \end{aligned} \tag{5}$$

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 $U_M\left(\theta, i, \mu_j\right)$:

$$2v - 2\left(\theta+1\right)^2 + \begin{cases} 4\theta\left(\alpha_i + \overline{\alpha}_i\left(\mu_j\right)\right) & \text{if } \theta \geqslant 0, \\ 4\theta\left(\alpha_i + \underline{\alpha}_i\left(\mu_j\right)\right) & \text{if } \theta < 0. \end{cases} \tag{6}$$

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

$$\begin{aligned} \overline{\alpha}_h\left(\mu_l\right) &= \frac{1}{2} + \frac{1}{2}\mu_l \beta\left(\gamma, \varepsilon\right), \\ \overline{\alpha}_l\left(\mu_h\right) &= \frac{1}{2} + \varepsilon + \frac{1}{2}\mu_h \beta\left(\gamma, \varepsilon\right), \end{aligned} \tag{7}$$

where

$$\beta\left(\gamma, \varepsilon\right) = \begin{cases} 2\gamma\left(\frac{1}{4} - \varepsilon^2\right) & \text{if } \gamma \leqslant \frac{2\varepsilon}{4\varepsilon^2+1}, \\ \gamma - \varepsilon & \text{if } \gamma \geq \frac{2\varepsilon}{4\varepsilon^2+1}. \end{cases} \tag{8}$$

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 (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 (7).

We observe from equation (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 find the distribution given prices, we need to know the gross 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 (6) and (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 2 (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. $U_M\left(\theta, i, \mu_j\right)$ *is supermodular in* θ *and* $\overline{\alpha}_i + \alpha_i$;
4. *for all* $i', j' \in \{l, h\}$, $j' \neq i'$ *and* $\theta \geqslant 0$, $U_M\left(\theta, i, \mu_j\right) \geqslant U_M\left(\theta, i', \mu_{j'}\right)$ *if and only if* $\alpha_i + \overline{\alpha}_i\left(\mu_j\right) \geqslant \alpha_{i'} + \overline{\alpha}_i\left(\mu_{j'}\right)$.

Notice that point (1) combined with equation (6) reveals the "product externality" effect. Point (2) helps us determine when the seller might create an unbalanced 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 seller creates a sorted or shuffled distribution.

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.

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.

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 (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

$$U_F(\theta, i) = U_M(\theta, i, 0). \tag{9}$$

Lemma 3 (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 2 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.

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. We can hence write the profits as a function of (μ_h, μ_l) as

$$\pi_M(\mu_h, \mu_l) = \mu_h(p_h(\boldsymbol{\mu}) - c) + \mu_l(p_l(\boldsymbol{\mu}) - c) \tag{10}$$

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

$$\frac{\partial \pi_M(\mu_h, \mu_l)}{\partial \mu_h} = \left[\mu_h \underset{(-)}{\frac{\partial p_h(\boldsymbol{\mu})}{\partial \mu_h}}\right] + \left[(p_h(\boldsymbol{\mu}) - c) + \mu_l \underset{(+)}{\frac{\partial p_l(\boldsymbol{\mu})}{\partial \mu_h}}\right] \tag{11}$$

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 (EQUILIBRIUM 1)
There exists a unique $0 < \rho_1 < \infty$ such that for $\rho \leqslant \rho_1$ the equilibrium is unique and is characterized by

1. *$p_i^* = \frac{1}{18} \max\left\{\frac{1}{2}\gamma\left(1 - 4\varepsilon^2\right), (\gamma - \varepsilon)\right\} + c$ for $i \in \{l, h\}$,*
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 $\left[\frac{1}{3}, 1\right]$.*

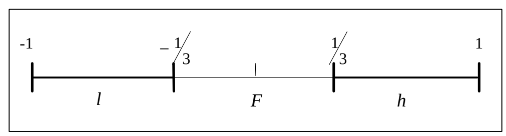

Figure 5: The equilibrium for $\rho \leqslant \rho_1$

Proposition 1 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 (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 (6) and (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, if ρ is low, the normalized utility of a buyer increases by very little as market share increases. Proposition 1 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.

Proposition 2 (EQUILIBRIUM 2)
There exists $\rho_1 < \rho_2 < \rho_3$ such that for $\rho_1 < \rho \leqslant \rho_3$ there exist 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 with $\mu_i^* < \mu_j^*$.*

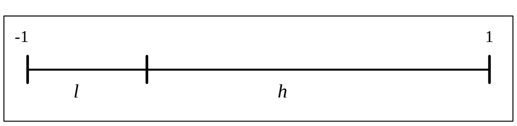

Figure 6: A symmetric equilibrium for $\rho_1 < \rho \leqslant \rho_3$

Proposition 2 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 and 2 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 3 (EQUILIBRIUM 3)
There exists $\rho_4 > \rho_3$ such that for $\rho_3 < \rho \leqslant \rho_4$ there exist 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 with $\mu_i^* < \mu_j^*$.*

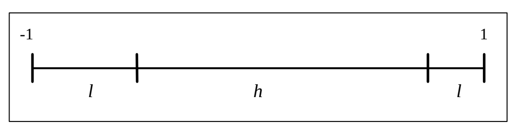

Figure 7: A symmetric equilibrium for $\rho_3 < \rho \leqslant \rho_4$

Proposition 3 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 4 (EQUILIBRIUM 4)
For $\rho > \rho_4$ there exist 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 and $\mu_i^* < \mu_j^*$.*

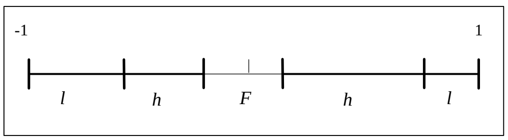

Figure 8: A symmetric equilibrium for $\rho > \rho_4$

After $\rho = \rho_4$, we see a reverse pattern. 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 $\rho > \rho_4$.

Proposition 5 summarize the comparative statics effects of an increase in the informativeness of γ.

Proposition 5 (COMPARATIVE STATIC)
In the perfect Bayesian equilibrium ,

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.*

5 Conclusion

We have shown that the existence of a recommender system creates additional surplus and introduces informational externalities into the pricing problem of the seller. If the output of the recommender system were independent of sales, then 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 charge a higher price for a higher quality product. Our findings show that when the recommender system does little to reduce the uncertainty, then this is indeed the way the seller handles 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.

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. A very simple way to think about this problem is to consider a world in which risk averse customers may not purchase the product because of the uncertainty about its true value. However, by offering information to the buyer which reduces the uncertainty, the seller could convince the buyer to purchase the product. 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. This may lead to the unbundling of recommendation and product and creates the possibility for the seller to charge for the recommendations separately. Finally, we did not consider uncertainty about one's own taste. With buyers to be uncertain about their tastes and the recommender system may be able to correlate their past purchases with aggregate information that would generate additional value to the customer.

Ultimately, we may be interested in the design of the recommendation mechanism itself. 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.

References

[1] Adams, W. J. and J. L Yellen (1976): "Commodity bundling and the burden of monopoly," *Quarterly Journal of Economics*, 90, 475-98.

[2] Ansari, A.,S. Essegaier, and R. Kohli (2000): "Internet Recommendation Systems", *Journal of Marketing Research,* 37, 363-375.

[3] Avery, C., P. Resnick, and R. Zeckhauser (1999): "The Market for Evaluations", *American Economic Review*, 89, 564-584.

[4] Avery, C., and R. Zeckhauser (1997), "Recommender systems for evaluating computer messages", *Communications of the ACM, 40, 88-89.*

[5] 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

[6] Brynjolfsson E. and M. D. Smith (2001): "The Great Equalizer? Consumer Choice Behavior at Internet Shopbots", *Sloan Working Paper 4208-01.*

[7] 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.

[8] Eppen, G. D., W. A. Hanson and R. K. Martin (1991): "Bundling-New Products, New Markets, Low Risk.", *Sloan Management Review,* 32, 7-14.

[9] IEEE Internet Computing: Industry Report: "Amazon.com Recommendations: Item-to-Item Collaborative Filtering", *http://dsonline.computer.org/0301/d/w1lind.htm.*

[10] 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

[11] Resnick, P. and H. Varian (1997): "Recommender Systems", *Communications of the ACM, 40, 56-58.*

[12] Schmalensee, R. L. (1984): "Gaussian demand and commodity bundling.", *Journal of Business*, 57, 211-230.