

1 Question

Suppose we have a web-platform that is designed to implicitly match customers and service providers via the past-user-based recommendation (examples of such a kind of platforms include *TripAdvisor*, *Yelp*, *etc*). The main goal of the platform is to render users satisfied with the quality of consumed products (in order to keep users using the platform). Users' reviews help the platform to infer about the qualities of different providers and based on this, to recommend users better options. However, once a new provider appears, the platform usually does not know the quality of the good provided, since there are no reviews yet. And the natural desire of the platform in this case is to figure out the quality of new alternative as quickly as possible in order to use (or not) it in recommendation for users. But how can the platform achieve this? Well, it seems that the platform can just start recommend the new alternative to users, collect reviews and then infer about the quality. But obviously the platform cannot do it too often and with too risky options : users will just stop using the platform that recommends them things that eventually turn out to be inferior. And here the first trade-off arises: how to induce exploration of new alternative without "sacrifice" customers too much? There are several papers on this topic that shed light on this trade-off and propose the optimal policy. The main idea in all those papers is that there is an optimal level of information obfuscation to create incentives for exploring new things.

However, in all previous works the quality of alternatives was exogenous to the process of recommendation. But in real life, providers are also strategic agents (and in most cases they are way more sophisticated than customers) who can change quality of their products in response to the recommendation policy the platform implements, and taking this into account the recommendation policy itself (I think) should change! So, questions I would be interested to address are: **what would be the optimal information revelation (recommendation) policy in this case, and how it is affected by the presence of strategic providers? Whether the quality of products increases (compare to the case when alternatives are exogenous)?**

With strategic alternatives the platform's and customers' preferences may become even more misaligned. Because now customers may be used by the platform not only for the sake of exploration (and to the benefit of subsequent agents) but also as a tool for encouraging the competition between alternatives.

2 Model

2.1 In words

We have a platform that is designed for providing customers with some services, and two firms (providers) are operating on this platform (it can be restaurants, hotels etc.).

Customers arrive sequentially and each should decide on the product from which firm to choose. Although the qualities of the products are unknown *ex ante*, customers have some priors and make their decision rationally taking into account all information available and updating their beliefs via the Bayes rule. After making a decision, each customer gets some reward, reports his experience to the platform (principal), and leaves the system.

The platform accumulates agents' reviews and based on this can recommend a product from a particular firm to the subsequent agents. The goal of the platform is to maximize the (infinite horizon ?) expected discounted reward of all agents.

Heterogeneous firms, in turn, are competing for agents and can invest to improve quality of their goods. The goal of each firm is to maximize its (infinite horizon ?) expected discounted profit taking into account the recommendation policy.

At $t = 0$ each firm privately learns its type. Then, the timing of the t -th ($t = 1, 2, \dots$) stage of the game is the following:

- 1) Each firm decides whether to stay in business or not (and incurs cost f_i if decides to stay)
- 2) Firms make private and costly investment in quality
- 3) The unobserved quality is realized
- 4) Customer arrives
- 5) The principal makes a recommendation
- 6) Customer decides which firm to choose
- 7) Payoffs are realized and the principal observes customers' experience

2.2 Formalization

2.2.1 Firms

We have two competing firms. Each firm has its own type that somehow is supposed to represent its "effectiveness" or "inclination to investing and sustaining the high quality of the good". The simplest example that comes in mind is that the type represents the firm's marginal cost of investment: the higher the type - the less effective the firm is. (types are either discrete θ^H, θ^L or continuous $\theta \in (0, 1)$).

Question: why do we need types?

Answer: customers will learn and infer about the type of the firm

Making investment $e \in \mathbb{R}_+$ (maybe make it discrete e^L and e^H) in quality of its product, each firm incurs cost $c_i(e)$ that should depend on the firm's type (for example $c_i(e) = \theta_i e_i$ where θ_i is the type of the i -th firm). Moreover, if the firm decides to stay in business, it incurs fixed cost f_i disregard whether the customer chooses the firm or not. f_i is low enough to ensure that at least one firm will operate in any case (maybe fixed cost also should dependent on type). Each firm obtains a profit 1 if the customer chooses its product. Denote for each firm

$$d_{it} = \begin{cases} 1, & \text{if consumer } t \text{ chooses the firm } i \\ 0, & \text{if consumer } t \text{ does not choose the firm } i \end{cases}$$

and

$$b_{it} = \begin{cases} 1, & \text{if the firm } i \text{ decided to stay in business at } t \\ 0, & \text{if the firm } i \text{ decided not to stay in business at } t \end{cases}$$

then each firm is maximizing

$$\max_{\{e_{it}, b_{it}\}_{t=1}^{\infty}} \mathbb{E} \left[\sum_{t=1}^{\infty} \delta^t b_{it} (d_{it} - c_i(e_{it}) - f_i) \right]$$

subject to the recommendation policy implemented by the principal.

Open questions:

can the firms observe the experience of users who chose the competitor? It seems reasonable if they could : the restaurant manager can send somebody to the nearby competitive restaurant to

infer about the real quality.

can the firm know that the competitor still in business?

Maybe infinite-horizon model will be too hard, make it finite (start from two-periods)

2.2.2 Products

Each firm provides a product which in time t has quality $q_{it} \in (0, 1)$ which is endogenously formed from the firms' investment e_{it} in this period.

Idea: the quality should somehow represent consumers' satisfaction of the product, the higher quality - the higher (expected) satisfaction.

Realization: Suppose that upon consuming the good from the firm i , the customer t gets a reward

$$r_{it} = \begin{cases} 1, & \text{w.p. } q_{it} \\ 0, & \text{w.p. } 1 - q_{it} \end{cases}$$

The quality of the product from the firm i at time t is determined as

$$q_{it} = \begin{cases} \frac{e_{it}}{e_{it} + e_{jt}}, & j \neq i, e_{it}^2 + e_{jt}^2 \neq 0 \\ 0, & e_{it} = e_{jt} = 0 \end{cases}$$

Question: why is the quality of the alternative i affected (negatively) by the investment made by the firm j ?

One possible answer: in real life there is no such a thing as an absolute quality. I think that good things are not good *per se*, they are good in comparison to other things. So, in this sense the assumption may seem to be reasonable.

Concern: in this set-up if both firms make high and equal investment the probability of success for the customer is $\frac{1}{2}$, whereas if only one firm made high investment while the other made low, the probability of success for the first firm becomes higher than $\frac{1}{2}$. Does it represent the reality?

2.2.3 Principal

Principal observes the history of consumers' choices and their payoffs, h^t (denote H^t the whole set of such histories) and can commit to the recommendation policy, a function $R : H^t \rightarrow M^t$, where M^t

is the message space feasible for the customer t . The principal's goal is to design a revelation policy $R(\cdot)$ in order to maximize agents' expected discounted payoff:

$$\pi = \max_{R(\cdot)} \mathbb{E} \left[\sum_{t=1}^{\infty} \delta^t r_t \right]$$

where r_t is an expected reward of the customer t .

2.2.4 Customers

Customers (agents) arrive sequentially one by one, they do not observe the experience of other agents, but can receive information revealed by the principal. For each principal's message $m_t \in M^t$ the agent's strategy is a function $\sigma^t : M^t \rightarrow A$, where $A = \{a_1, a_2\}$ describes the set of possible alternatives (products from the firm 1 or 2). In other words, the agent's strategy prescribes for each principal's message an action, the product from which firm to choose. Choosing the strategy, each agent t maximizes

$$\max_{\{a_1, a_2\}} \mathbb{E}_t [q_{it} | m_t]$$

Questions: what are priors? which alternative do the customers choose when indifferent between two?

3 Analysis

3.1 Straightforward observations

First of all, it seems that in this model it will never be optimal (starting from some point) to start recommending only one firm. This is because in this case the firm which is not recommended leaves the business (because of fixed cost) and this eliminates any incentive for the incumbent to invest. Hence, the quality becomes 0 and all subsequent customers will get 0. So, the platform should internalize its role in sustaining a competition.

3.2 No disclosure at all

Suppose that the platform employs the non-disclosure policy. That is, agents do not possess any information regarding the previous users' experience. In this case, agents make decisions based on

priors. If customers are homogeneous and have identical priors about the quality, then the firm which the customers are *a priori* inclined towards stays in business with zero investment while the second firm exits the market. In this equilibrium all customers get zero, so the platform's payoff is zero.

As we can see, now with strategic agents the non-disclosure policy becomes much worse. One reason is that it completely eliminates any incentive to increase the quality of alternatives.

3.3 Full disclosure

3.3.1 1 period, observable qualities

Suppose we have only 1 period, at the beginning of the period, firms privately draw its types θ according to some full-support continuous distribution with the p.d.f. $f(\theta)$. Then both firms simultaneously perform investment. After that the customer arrives and can observe realized qualities. The customer chooses the firm with higher quality (ties do not happen almost surely). In this case the following Lemma is true

Lemma 1. *Suppose $\theta \sim U(0,1)$ then in the one period simultaneous-move game with observable qualities there exist a symmetric equilibria in which both firms invest $e(\theta) = -\ln(\theta)$.*

Proof. See Appendix □

3.3.2 Many periods, observable qualities

If there are more than 1 period then in all periods after the first one the firm with higher type will learn that its type is high (because the customer did not choose it in the first period) and that is why the aforementioned equilibrium breaks down: less effective firm is better-off if it leaves the business, but if one of the firm leaves then the other has no incentive to invest: it becomes a monopoly. So let us note this result

Lemma 2. *Suppose $\theta \sim U(0,1)$ then in many-periods simultaneous-move game with observable qualities there exists an equilibrium in which:*

1. *In the first period both firms play $e(\theta) = -\ln(\theta)$*
2. *In all subsequent periods:*

- the more efficient firm (with smaller θ) plays $e(\theta) = -\ln(\theta)$ if and only if there is a competitor in business, and $e(\theta) = 0$ otherwise
- the less efficient firm (with higher θ) leaves the business forever.

Proof. See Appendix. □

As we can see, if qualities are observable in many periods game, unravelling may happen, and the principal's expected payoff becomes equal to the expected reward of the first agent

$$\pi_0 = r_1 = \mathbb{E} \left[\frac{\ln \theta_i}{\ln \theta_i + \ln \theta_j} \middle| \theta_i < \theta_j \right] = \frac{3}{4}, \quad i \neq j$$

3.4 Not full disclosure

3.4.1 Qualities are observable by the principal

Can the principal do better? Let us take the simplest case, when the principal can observe the qualities. Suppose the principal commits to recommend high quality firm with probability q and low probability good with probability $1 - q$. In this case it is easy to see that the symmetric equilibrium of one-period game is an equilibrium in many-periods game

Lemma 3. *Suppose $\theta \sim U(0, 1)$, $f \leq 0.5$, the platform can observe the realized qualities and commits to recommend to the customer better option with probability $q \in (0.5, 1)$ and worse option with probability $1 - q$. Then there exists a symmetric equilibrium in which both firms are playing $e(\theta) = -(2q - 1) \ln \theta$.*

Moreover, in this equilibrium the platform's expected payoff is strictly greater than the platform's expected payoff π_0 in case of unravelling (Lemma 2)

Proof. See Appendix. □

As we can see, the platform can encourage competition.

4 Extensions

1. Agents do not know they place in line.
2. Heterogeneous customers

Appendix

Proof of Lemma 1. Let us find the equilibrium strategy for firms in a form $\varepsilon(\theta)$ where $\varepsilon(\theta) : (0, 1) \rightarrow \mathbb{R}$ is a decreasing function. Then each firm is maximizing

$$\max_e \mathbb{P}(\theta > \varepsilon^{-1}(e))(1 - \theta e - f) - \mathbb{P}(\theta \leq \varepsilon^{-1}(e))(\theta e + f)$$

or using the fact that

$$\mathbb{P}(\theta \leq x) = \begin{cases} 1, & x > 1 \\ x, & x \in [0, 1] \\ 0, & x < 0 \end{cases}$$

the firm is maximizing

$$\max_e 1 - \varepsilon^{-1}(e) - \theta e - f$$

F.O.C.:

$$-\frac{1}{\varepsilon'(\varepsilon^{-1}(e))} = \theta$$

Since the equilibrium we are interested in is symmetric, $\varepsilon(\theta) = e$ and hence

$$\varepsilon'(\theta) = -\frac{1}{\theta}$$

Assuming $\varepsilon(1) = 0$ we get

$$\varepsilon(\theta) = -\ln \theta$$

□

Proof of Lemma 3. If both firms stay in business in the particular period, then let us find the equilibrium strategy for firms in a form $\varepsilon(\theta)$ where $\varepsilon(\theta) : (0, 1) \rightarrow \mathbb{R}$ is a decreasing function. Then each firm is maximizing

$$\begin{aligned} & \max_e q \{ \mathbb{P}(\theta > \varepsilon^{-1}(e))(1 - \theta e - f) + \mathbb{P}(\theta \leq \varepsilon^{-1}(e))(-\theta e - f) \} + \\ & + (1 - q) \{ \mathbb{P}(\theta > \varepsilon^{-1}(e))(-\theta e - f) + \mathbb{P}(\theta \leq \varepsilon^{-1}(e))(1 - \theta e - f) \} \end{aligned}$$

or using the fact that

$$\mathbb{P}(\theta \leq x) = \begin{cases} 1, & x > 1 \\ x, & x \in [0, 1] \\ 0, & x < 0 \end{cases}$$

the firm is maximizing

$$\max_e q(1 - \varepsilon^{-1}(e)) + (1 - q)\varepsilon^{-1}(e) - \theta e - f$$

F.O.C.:

$$-\frac{q}{\varepsilon'(\varepsilon^{-1}(e))} + \frac{1 - q}{\varepsilon'(\varepsilon^{-1}(e))} - \theta = 0$$

Since the equilibrium we are interested in is symmetric, $\varepsilon(\theta) = e$ and hence

$$\varepsilon'(\theta) = -\frac{2q - 1}{\theta}$$

Assuming $\varepsilon(1) = 0$ we get

$$\varepsilon(\theta) = -(2q - 1) \ln \theta$$

It remains to check that less effective firm stays in business. Expected profit from staying in business for less effective firm is equal to

$$\mathbb{E}[1 - q + (2q - 1)\theta \ln \theta - f] = \frac{5}{4} - \frac{3}{2}q - f \geq 0$$

The expected payoff of the platform is equal to

$$\pi = \frac{q}{1 - \delta} \mathbb{E} \left[\frac{\ln \theta_i}{\ln \theta_i + \ln \theta_j} \middle| \theta_i < \theta_j \right] + \frac{1 - q}{1 - \delta} \mathbb{E} \left[\frac{\ln \theta_j}{\ln \theta_i + \ln \theta_j} \middle| \theta_j \geq \theta_i \right] = \frac{2q + 1}{4(1 - \delta)} > \frac{3}{4} = \pi_0$$

□