Nonlinear Pricing with Network Externalities and Countervailing Incentives

Meng, Dawen and Tian, Guoqiang

January 2008

Nonlinear Pricing with Network Externalities and Countervailing Incentives*

Dawen Meng †
International Business School
Shanghai Institute of Foreign Trade
Shanghai, China. 201600

Guoqiang Tian ‡
Department of Economics
Texas A&M University
College Station, Texas 77843

January, 2008/revised: August, 2008

Abstract

This paper considers the screening problem faced by a monopolist of a network good in a general setting. We demonstrate that the joint presence of asymmetric information and network externalities revise the "no distortion on the top" and "one-way distortion" principle. The pattern of consumption distortions crucially depends on the congestion of the network. It exhibits one-way distortion in un-congestible network and two-way distortion in congestible network. The countervailing incentives problem from potential entry threat is also analyzed. As the competitiveness of the outside competitors increases, the incumbent firm should adjust its nonlinear pricing scheme accordingly, which will distort the allocations of both types.

Keywords: Nonlinear pricing, network externalities, countervailing incentives, type-dependent reservation utility

Journal of Economic Literature Classification Number: D42, D62, D82

1 Introduction

Two prevailing rules in the classical adverse selection model are "*no distortion on the top*" and "*one-way distortions for less efficient types*", which means under asymmetric information, the most efficient agent has the first-best allocation, all the other types have allocations lower than the first-best one. These rules were supported by many seminal studies such as those in

*The second author thanks the National Natural Science Foundation of China (NSFC-70773073) and Private Enterprise Research Center at Texas A&M University for Financial support.

†Email: devinmeng@yahoo.com

‡Email:tian@econmail.tamu.edu

Maskin and Riley (1984), Baron and Myerson (1982), Mussa and Rosen (1978), Myerson (1981), etc. However, these rules may not be true in the presence of *externalities* or in the case where the agents face *countervailing incentives.* In this paper, we will discuss both the cases in a nonlinear pricing environment.

An externality is present whenever the well-being of a consumer or the production possibilities of a firm are directly affected by the actions of another in the economy. In consumption sectors, externalities arise whenever a person's utility function includes arguments controlled by the other consumers. Among all forms of consumption externalities, network externality is the most typical one. It is defined as a change in the benefit, or surplus, that an agent derives from a good when the total consumption of the same kind of goods changes. Such a phenomenon might arise for any of a number of reasons: because the usefulness of the product depends directly on the size of the network (e.g., telephones, fax machines); or because of bandwagon effect, which means the desire to be in style: to have a good because almost everyone else has it; or indirectly through the availability of complementary goods and services (often known as the "hardware-software paradigm") or of postpurchase services (e.g., for automobiles).

Most studies on consumption externalities in principal-agent setup find that "no distortion on the top" rule does not hold any more, but the consumptions still exhibit "one-way distortion". Hahn (2003) built a model of telecommunication to examine the role of call and network externalities in nonlinear pricing. He concludes that in equilibrium all types end up with sub-optimal quantities. And he attributes this result to the existence of call externalities which is the benefit of incoming calls to a subscriber who does not have to pay for the calls. Segal (1999, 2003) developed a general model of contracting with externalities and characterize the nature of the arising inefficiencies. When externalities are positive, he shows that each agent's consumption level is smaller in the resulting equilibrium allocation than in the socially efficient one. Strategic complementarity is identified as the factor accounting for this result. However, his analysis makes two additional assumptions that are useful in identifying the direction of distortions: first, the consumers are identical, and second, total welfare depends only on aggregate trade and not on its allocation across consumers. Csorba (2008) showed that under-consumption result holds even without these two assumptions if externalities are positive. Still based on the critical assumption of strategic complementarity, he builds a general model to describe and solve the screening problem faced by the monopolist of a network good. Applying monotone comparative static tools, he demonstrates that the joint presence of asymmetric information and positive network effects leads to a strict downward distortion for all consumers in the quantities provided.

Another revise to the "*no distortion on the top*" and "*one-way distortion*" rules arises when agents face *countervailing incentives*. When the reservation utility of an agent, as well as his cost of acting for the principal, depend on his private information, the low-type agents may have incentives to choose the contract menu of the high-type agents. Lewis and Sappington (1989) termed this phenomenon "countervailing incentives" and give a pioneering study on this issue. They showed when the fixed cost of a regulated firm is type-dependent, its payoff function is not monotone on the whole interval of types, and the agent can be tempted to overstate his private information for some of his realizations and to understate it for others. The performance can be distorted both above and below the efficient level. Maggi and Rodriguee-Clare (1995) gave a complete analysis of the principal-agent problem with countervailing incentives. They show that the pattern of the two-way distortion depends crucially on whether the reservation utility of the agent is convex or concave in his private information. If it is concave or mildly convex function, then the low-type agent's output is distorted downward, and high-type agent's output is distorted upward. If the reservation utility is strongly convex, then the opposite is the case.

This paper analyzes the nonlinear pricing problem faced by an entry-deterring monopolist seller under asymmetric information and network externalities. We first show that the one-way distortion results in Segal (1999, 2003), and Csorba (2008) depend on the complementarity between consumers. If the consumers are rivals or substitutes to each other, the one-way distortion results are not true in general. In the presence of network externalities, the congestion of network accounts for the substitutability between consumers. All the participants in a network interlinked with each other in two aspects: first, everyone benefits from others' consumption; second, they are rivals to each other in the process of competing for limited network resource. For example, all the internet subscriber benefit from increasing number of new subscribers with whom they can share information with. One the other hand, they may also suffer from network jam because of bandwidth limit. If the former effect dominates, the network is "un-congestible". If the latter one dominate, it is "congestible". In this paper, we give a full characterization of the pattern of distortions in the presence of network externalities. We show that the one-way distortion attained by Segal (1999, 2003) and Csorba (2008) presents if the network is un-congestible. However, the consumption exhibits two-way distortion if the network is congestible.[1]

[1]Lockwood (2003) studied the production externalities problem in a principal-agent model and a similar two-way distortion result. In his model, the cost of any agent depends positively on the output offered by the other agents. It is shown that the optimal contract for the principal exhibits two-way distortion: the output of any agent is oversupplied (relative to the first-best) when his marginal cost of effort is low, and undersupplied when

We also consider the distortion way under co-existence of the network externalities and the countervailing incentives problems. Although both factors can revise the "no distortion on the top" and "one-way distortion" rule, the joint presence of them makes the distortion pattern in the nonlinear pricing contract more complicated, which, to our knowledge, has been little studied in the existing literature. Facing the outside market, consumers can choose to bypass the network of incumbent firm and switch to buying from outside competitors, it results in type-dependent reservation utilities of consumers, which is the causality of countervailing incentives problem.

In the classical countervailing incentives research such as Lewis and Sappington (1989) and Maggi and Rodriguee-Clare (1995), the nature of optimal contract depends crucially on the shape of reservation utility function. In this paper, we show that the shape of reservation utility function is in turn determined by the marginal cost of outside competitors. As such, the incumbent firm should adjust its nonlinear pricing contract according to the cost of potential entrants, which means that competitors with different competitiveness pose different degree of threat to the incumbent firm. Laffont and Tirole (1990) also discussed the problem of optimal network bypassing in a natural monopolist regulation economic environment. In the presence of bypass, the monopolist may have to offer advantageous term to high-demand consumers in order to retain them. This may lead low-demand consumers to consume high-demanded consumers' bundle. They discussed the "upward-binding" incentive constraints and countervailing incentives problem. However, in Laffont and Tirole (1990), the good displays no network externalities although the regulated firm is in network industry. There is thus no interaction between different types' consumption. While in our model, the consumptions of different types are "interdependent" through the network.

The reminder of this paper is organized as follows: Section 2 sets up economic environments and gives the benchmark result without network externalities. Section 3 considers the design of optimal nonlinear pricing contract in the presence of network externalities. Section 4 discusses network bypassing and countervailing incentives problem. Finally, concluding remarks are offered in Section 5.

his marginal cost of effort is high. A critical assumption underling this result requires a non-increasing marginal cost of effort which is satisfied when agents operate production process with non-decreasing return to scale. This condition is analogous to the "substitutability" or "congestion" condition in our model.

2 Economic environments

Consider a principal-agent model in which the principal is a monopolist of a network good with marginal production cost c and total output q. The principal's payoff function is given by $V = t - cq$, where t is the payment received from consumers.

Consumers have heterogenous preferences for the good, suppose there are n different types of consumers, so that $\theta_1 < \theta_2 < \cdots < \theta_n$ and let Θ denote the set of different types. Let $f(\theta_i)$ represents the proportion of θ_i type consumers in the whole population. Denoted by $\Delta\theta \equiv \theta_i - \theta_{i-1} > 0$ the difference of adjacent types and $F(\theta_i) = \sum_{j \leqslant i} f(\theta_j)$, the cumulative frequency. Then by Law of Large Numbers, it is mathematically equivalent to the framework with a single agent whose type θ is distributed with c.d.f $F(\cdot)$ and p.d.f $f(\cdot)$ on $\Theta = \{\theta_1, \cdots, \theta_n\}$.

A consumer of θ_i type is assumed to have an utility function of $U_i = \theta_i V(q_i) + \Psi(Q) - t_i$, where q_i is the amount of the network good he consumes, $Q = \sum_i f(\theta_i) q_i$ is the total amount of network good in the economy (network size) and t_i is the tariff charged for q_i by the principal. The utility function can be divided into two terms addictively: $\theta_i V(q_i)$ is the *intrinsic value* of consuming, while $\Psi(Q)$ is the *network value*. Note that, we assume the network effect is homogeneous among all the consumers, namely, the network value is independent of individual preference θ_i and individual consumption q_i.

It is assumed that $V'(q) > 0$ and $V''(q) < 0$. The congestion of network is defined by the sign of $\Psi''(Q)$.

Definition 1 *The network is congestible if and only if* $\Psi''(Q) < 0$.

Remark 1 If the network capacity is large and the maintaining technology is advanced enough, an increase in one consumer's consumption will increase the marginal utilities of others, and so $\Psi''(Q) > 0$. On the contrary, if network capacity and maintaining technology are limited, consumers are rivals to each other in the sense that an increase in one consumer's consumption will decrease the marginal utilities of others, and thus $\Psi''(Q) < 0$.

The objective of the monopolist is to design a menu of incentive-compatible and self-selecting quantity-price pairs $\{q(\hat{\theta}), t(\hat{\theta})\}$ to maximize his own revenue, where $\hat{\theta} \in \Theta$ is the consumer's announcement. The timing of contracting is given by the following procedures:

- Stage 1. The consumers observe their own "type" θ.
- Stage 2. The monopolist announces their quantity and pricing schedule $\{q(\hat{\theta}), t(\hat{\theta})\}$.

- Stage 3. Based on her type θ, each consumer determines her optimal announcement: $\hat{\theta} = \arg\max_{\tau}\{\theta V(q(\tau)) + \Psi(Q) - t(\tau)\}$ subject to her individual rationality constraint (IR): $\theta V(q) + \Psi(Q) - t \geqslant \hat{U}(\theta)$, where $\hat{U}(\theta)$ is her reservation utility.[2]

- Stage 4. The monopolist gets payoff of $E_{\theta\in\Theta}\big[t(\theta) - cq(\theta)\big]$ and each consumer gets her information rent.

If the monopolist offers the same menu of contracts for all consumers, a consumer may have an incentive to mimic the other types. As in the standard adverse selection model, an incentive-compatible menu $\{q_i, t_i\}_{i=1}^{n}$ should satisfy participation constraints (IR_i) and incentives constraints (IC_{ij}):

$$\begin{aligned} IR_i &: \theta_i V(q_i) + \Psi(Q) - t_i \geqslant 0 \\ IC_{ij} &: \theta_i V(q_i) + \Psi(Q) - t_i \geqslant \theta_i V(q_j) + \Psi(Q) - t_j \end{aligned}$$

or equivalently, written in the form of information rents as:

$$\begin{aligned} IR_i &: U_i \geqslant 0 \\ IC_{ij} &: U_i \geqslant U_j + (\theta_i - \theta_j)V(q_j), \forall i, j, \end{aligned}$$

where the reservation utilities of the consumers are normalized to be zero.

In the absence of network externalities ($\Psi(Q) = 0$), the standard problem of the monopolist is:

$$\begin{cases} \max\limits_{\{t_i,q_i\}} \left\{ \sum\limits_{i=1}^{n} f(\theta_i)\,(t_i - cq_i) \right\} \\ s.t.\ IR_i : \theta_i V(q_i) - t_i \geqslant 0 \\ \qquad IC_{ij} : \theta_i V(q_i) - t_i \geqslant \theta_i V(q_j) - t_j, \forall i, j \end{cases}$$

or equivalently,

$$\text{(P1)}\begin{cases} \max\limits_{\{U_i,q_i\}} \left\{ \sum\limits_{i=1}^{n} f(\theta_i)\big[\theta_i V(q_i) - cq_i\big] - \sum\limits_{i=1}^{n} f(\theta_i) U_i \right\} \\ s.t.\ IR_i : U_i \geqslant 0 \\ \qquad IC_{ij} : U_i \geqslant U_j + (\theta_i - \theta_j)V(q_j), \forall i, j \end{cases}$$

We first analyze the set of constraints to find the binding ones.

[2] Note that, because of a continuum of consumers, the individual misreport has no effect on the total consumption Q.

Lemma 1 *In the second-best optimal contract there are n binding constraints: the individual rationality constraint of the lowest-type consumer is binding, and the incentive constraints are binding for $i = 2, 3, \ldots, n$.*

Proof. See Appendix. ■

Applying Lemma 1, in the absence of network externalities, we get the benchmark MR (Maskin-Riley) contract from (P1):[3]

$$\left[\theta_i - H(\theta_i)\right] V'(q_i^{MR}) = c, \forall i \in \{1, 2, \cdots, n\} \tag{1}$$

and

$$U_1^{MR} = 0, U_i^{MR} = \sum_{j=1}^{i-1} \Delta\theta V(q_j^{MR}), \forall i \geqslant 2 \tag{2}$$

where $H(\theta) = \frac{1-F(\theta)}{f(\theta)}\Delta\theta$, and $\frac{1-F(\theta)}{f(\theta)}$ is the reciprocal hazard rate of distribution $F(\theta)$.

It is easy to check that if the monotone hazard rate condition $H'(\theta) \leqslant 0$ is satisfied, $q_i^{MR} < q_{i+1}^{MR}; q_i^{MR} < q_i^*, \forall i < n$ and $q_n^{MR} = q_n^*$, where q_i^* are the first-best outputs that are obtained from $\theta_i V'(q_i^*) = c$. It is clear that under asymmetric information the consumptions of all types except the one with highest willing to pay will be distorted downward, which is consistent with the rules "no distortion on the top" and "one-way distortion" .

3 Economies with network externalities

3.1 Two-type case

In this subsection, we assume that $\Theta = \{\underline{\theta}, \overline{\theta}\}, \Pr(\theta = \underline{\theta}) = v, \Pr(\theta = \overline{\theta}) = 1 - v$, the network magnitude is thus $Q = v\underline{q} + (1-v)\overline{q}$. Under complete information, the monopolist's problem is:

$$\begin{cases} \max\limits_{\{(\underline{t},\underline{q});(\overline{t},\overline{q})\}} & v(\underline{t} - c\underline{q}) + (1-v)(\overline{t} - c\overline{q}) \\ s.t.\ IR(\underline{\theta}) : & \underline{\theta}V(\underline{q}) + \Psi(Q) - \underline{t} \geqslant 0 \\ \quad IR(\overline{\theta}) : & \overline{\theta}V(\overline{q}) + \Psi(Q) - \overline{t} \geqslant 0. \end{cases}$$

or equivalently,

$$\text{(P2)}\begin{cases} \max\limits_{\{(\underline{U},\underline{q});(\overline{U},\overline{q})\}} v\left[\underline{\theta}V(\underline{q}) - c\underline{q}\right] + (1-v)\left[\overline{\theta}V(\overline{q}) - c\overline{q}\right] + \Psi(Q) - \left[v\underline{U} + (1-v)\overline{U}\right] \\ s.t.\ IR(\underline{\theta}) : \underline{U} \geqslant 0 \\ \quad IR(\overline{\theta}) : \overline{U} \geqslant 0 \end{cases}$$

[3] It is assumed that $\theta_i - H(\theta_i) > 0, \forall i \in \{1, 2, \cdots, n\}$ to guarantee that $q_i > 0$.

The first-best consumption is thus:

$$\begin{cases} \underline{\theta} V'(\underline{q}^{FB}) + \Psi'\left(v\underline{q}^{FB} + (1-v)\overline{q}^{FB}\right) = c, \\ \overline{\theta} V'(\overline{q}^{FB}) + \Psi'\left(v\underline{q}^{FB} + (1-v)\overline{q}^{FB}\right) = c. \end{cases} \tag{3}$$

Under asymmetric information, two incentive compatible constraints should be added to the above program, then we get:

$$(\text{P3}) \begin{cases} \max\limits_{\{(\underline{U},\underline{q});(\overline{U},\overline{q})\}} v\left[\underline{\theta} V(\underline{q}) - c\underline{q}\right] + (1-v)\left[\overline{\theta} V(\overline{q}) - c\overline{q}\right] + \Psi(Q) - \left[v\underline{U} + (1-v)\overline{U}\right] \\ s.t.\ IR(\underline{\theta}) : \underline{U} \geqslant 0 \\ \qquad IR(\overline{\theta}) : \overline{U} \geqslant 0 \\ \qquad IC(\underline{\theta}) : \underline{U} \geqslant \overline{U} - \Delta\theta V(\overline{q}) \\ \qquad IC(\overline{\theta}) : \overline{U} \geqslant \underline{U} + \Delta\theta V(\underline{q}) \end{cases}$$

From Lemma 1, the downward incentive constraint $IC(\overline{\theta})$ and the participation constraint of the low-demand type $IR(\underline{\theta})$ are binding, then the consumptions in the second-best contract are characterized by the following first-order conditions:

$$\begin{cases} \left(\underline{\theta} - \dfrac{1-v}{v}\Delta\theta\right) V'(\underline{q}^{SB}) + \Psi'\left(v\underline{q}^{SB} + (1-v)\overline{q}^{SB}\right) = c, \\ \overline{\theta} V'(\overline{q}^{SB}) + \Psi'\left(v\underline{q}^{SB} + (1-v)\overline{q}^{SB}\right) = c. \end{cases} \tag{4}$$

We synthesize the first-best and second-best solution by considering them as solution to the following parameterized form:

$$\max_{\{\underline{q},\overline{q}\}} \Pi(\underline{q}, \overline{q}, \alpha) \tag{5}$$

where

$$\Pi(\underline{q}, \overline{q}, \alpha) = v\left[\alpha V(\underline{q}) - c\underline{q}\right] + (1-v)\left[\overline{\theta} V(\overline{q}) - c\overline{q}\right] + \Psi(Q).$$

Note that we have the first-best contract given in (3) when $\alpha = \underline{\theta}$, and the second-best contract given in (4) when $\alpha = \underline{\theta} - \frac{1-v}{v}\Delta\theta$.

We then have the following proposition.

Proposition 1 *In the presence of network externalities and asymmetric information, if $V'(q) >$, $V''(q) < 0$, the Hessian matrix Π_{qq} is negative definite at the solution to (5) for all $\alpha \in [\underline{\theta} - \frac{1-v}{v}\Delta\theta, \underline{\theta}]$, the direction of distortion in consumptions depends on whether the network is congestible, namely, the sign of $\Psi''(Q)$.*

1. *If $\Psi''(Q) > 0$, $\underline{q}^{SB} < \underline{q}^{FB}$ and $\overline{q}^{SB} < \overline{q}^{FB}$. Consequently, the consumption exhibits one way distortion: the second-best consumption is smaller than the first-best one for all types.*

2. *If $\Psi''(Q) < 0$, $\underline{q}^{SB} < \underline{q}^{FB}$ and $\overline{q}^{SB} > \overline{q}^{FB}$. Consequently, the consumption exhibits two-way distortion: the consumption of the low demand consumers is distorted downward, while that of the high demand consumers is distorted upward.*

For both cases, the network magnitude is downsized: $Q^{SB} < Q^{FB}$.

Proof. See Appendix. ■

The result can be interpreted as follows. In order to minimize the information rent captured by the high-type consumer, the allocation of low-type agent should be lowered relative to the first-best level. This is the basic trade off between allocation efficiency and rent extraction in the standard adverse selection settings. In contrast to the canonical setting, the consumptions of different consumers in our model interact through network effects. If the network is un-congestible, they are complementary in the sense that the consumptions of high type agent is also distorted downward. On the contrary, if the network is congestible, the consumers in the network are rivalry, then the consumption of the high-type consumers is distorted upward to counteract the decrease in that of the low type.

3.2 Multi-type case

In this subsection we extend the model to economic environments with more than two types of agents. Under complete information, similar to the case of two types, the first-best consumptions are:

$$\theta_i V'(q_i) + \Psi'(Q) = c \ \forall i \in \{1, 2, \cdots, n\}. \tag{6}$$

Under asymmetric information, with the incentive compatible constraints and participation constraints, the monopolist's optimization program can be represented as:

$$(\text{P4}) \begin{cases} \max\limits_{\{U_i, q_i\}} \left\{ \sum\limits_{i=1}^{n} f(\theta_i) \left[\theta_i V(q_i) - c q_i \right] + \Psi\left(\sum\limits_{i=1}^{n} f(\theta_i) q_i \right) - \sum\limits_{i=1}^{n} f(\theta_i) U_i \right\} \\ s.t. \ IR_i : U_i \geqslant 0, \\ \quad IC_{ij} : U_i \geqslant U_j + (\theta_i - \theta_j) V(q_j), \forall i, j. \end{cases}$$

From Lemma 1 we have $U_i = \Delta\theta \sum_{j=1}^{i-1} V(q_j)$. Then the principal's objective function in (P4) can be rewritten as:

$$\sum_{i=1}^{n} \left[\theta_i - \frac{1 - F(\theta_i)}{f(\theta_i)} \Delta\theta \right] f(\theta_i) V(q_i) + \Psi\left(\sum_{i=1}^{n} f(\theta_i) q_i \right) - c \sum_{i=1}^{n} f(\theta_i) q_i \tag{7}$$

The first-best and second-best consumptions can thus be regarded as the solutions to the following unconstrained program parameterized on ϵ:

$$\max_{\mathbb{q} \in \mathbb{R}_+^n} \Pi(\mathbb{q}, \epsilon) \tag{8}$$

where

$$\Pi(\mathbb{q}, \epsilon) = \sum_{i=1}^{n} \Big[\theta_i + \epsilon H(\theta_i)\Big] f(\theta_i) V(q_i) + \Psi\left(\sum_{i=1}^{n} f(\theta_i) q_i\right) - c \sum_{i=1}^{n} f(\theta_i) q_i$$

with $\mathbb{q} = (q_1, q_2, \cdots, q_n) \in \mathbb{R}_+^n$ and $\epsilon \in [-1, 0]$. If $\epsilon = 0$ the solution to (8) is the first-best contract; if $\epsilon = -1$ it is the second-best contract characterized by the following proposition.

Proposition 2 *If the weak monotone hazard rate condition* $\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\frac{1-F(\theta)}{f(\theta)}\right] \leqslant 0$ *is satisfied, and the Hessian matrix* Π_{qq} *is negative definite at* $\left(q_1^{SB}, q_2^{SB}, \cdots, q_n^{SB}\right)$ *for* $\epsilon = -1$*, then the second-best consumptions* q_i^{SB} *satisfy*

$$\Big[\theta_i - H(\theta_i)\Big] V'(q_i) + \Psi'\left(\sum_{i=1}^{n} f(\theta_i) q_i\right) = c, \forall i \in \{1, 2, \cdots, n\} \tag{9}$$

the second-best information rents are

$$U_1^{SB} = 0, U_i^{SB} = \Delta\theta \sum_{j=1}^{i-1} V(q_j^{SB}), \forall i \in \{2, 3, \cdots, n\} \tag{10}$$

the second-best tariffs are

$$\begin{aligned} t_1 &= \theta_1 V(q_1^{SB}) + \Psi\left(\sum_{i=1}^{n} f(\theta_i) q_i^{SB}\right); \\ t_i &= \theta_i V(q_i^{SB}) - \Delta\theta \sum_{j=1}^{i-1} V(q_j^{SB}) + \Psi\left(\sum_{i=1}^{n} f(\theta_i) q_i^{SB})\right), \forall i \in \{2, 3, \cdots, n\}. \end{aligned} \tag{11}$$

Proof. See Appendix. ■

Similarly, we now give the following proposition.

Proposition 3 *Suppose that* $V'(\cdot) > 0$, $V''(\cdot) < 0$*, the weakly monotone hazard rate property* $\frac{\mathrm{d}}{\mathrm{d}\theta}\left[\frac{1-F(\theta)}{f(\theta)}\right] \leqslant 0$ *holds, and Hessian matrix* Π_{qq} *is negative definite for all* $\mathbb{q} \in \mathbb{R}_+^n$ *and* $\epsilon \in [-1, 0]$*. Then distortion way of consumptions depends on the sign of* $\Psi''(Q)$*.*

1. *If the network is un-congestible,* $\Psi''(Q) > 0$*, then the consumption exhibits one-way distortion:* $q_i^{SB} < q_i^{FB}, \forall i$

2. *If the network is congestible,* $\Psi''(Q) < 0$*, it exhibits two-way distortion. That means there exists a threshold value* $i^* \in \{1, 2, \cdots, n\}$*, for* $i > i^*$*,* $q_i^{SB} > q_i^{FB}$*; for* $i < i^*$*,* $q_i^{SB} < q_i^{FB}$*.*

In both cases the network magnitude will be downsized: $Q^{SB} < Q^{FB}$*.*

Proof. See Appendix. ■

The first conclusion of Proposition 3 is in line with the main conclusion in Hahn (2003), Segal (1999, 2003), and Csorba (2008). $\Psi''(\cdot) > 0$ implies that the marginal value from an increase in individual consumption is higher at a higher level of others' consumption : $\frac{\partial^2 U_i}{\partial q_i \partial q_j} > 0, \forall i \neq j$. Its interpretation is that the externalities in bigger network is larger than in small network, so an agent is more eager to consume more when other agent consume more. It is consistent with the critical "strategic conplementarity" assumption in Segal (1999,2003) and Csorba (2008). This condition allows us to characterize the optimal contracts in a general setting by applying monotone comparative static tools, pioneered by Topkis (1978) and Milgrom and Shannon (1994) without solving equation (21) in appendix D explicitly.[4] The second conclusion of the above proposition is new. For every type θ_i, the principal has incentive to lower his consumption to reduce the information rent accrued to more efficient agents with type $\theta > \theta_i$; on the other hand, the principal also has an incentive to raise every type's consumption to increase the whole network value. If the network is un-congestible, the rent-reducing incentive dominates for all types, so everyone's consumptions is lower than the first-best case; if the network is congestible, for the high type consumers, the network- value-increasing incentive dominates, while for the low type consumers the rent-reducing incentive dominates, so two-way distortion is the outcome.

4 Entry Deterrence in Congestible Network: The Countervailing Incentives Problem

In this section we discuss another causality for the failure of "no distortion on the top" and "one-way distortion" rules. That is the countervailing incentive problem faced by agents. We assume that the consumers can bypass the network offered by the incumbent firm and enter a competitive market including many homogenous firms. All these firms are the potential entrants of the market. Let ω denote the marginal production cost of the entrants. We assume that the goods or services offered by the entrants are incompatible with that of the incumbent monopolist, [5] and they have not yet formed their own consumers network. In the competitive outside market, each firm's unit charge equals its marginal cost ω, so the representative consumer's utility derived from consuming the entrants' goods is $U(q) = \theta V(q) - \omega q$. Let $G(\theta) = \max_q [\theta V(q) - \omega q]$. For simplicity, we discuss the two-type case in this section, define $\underline{G} = G(\underline{\theta})$, $\overline{G} = G(\overline{\theta})$, and $\Delta G = \overline{G} - \underline{G}$.

[4] By Topkis (1978) and Milgrom and Shannon (1994), a twice continuously differentiable function $\Pi = \Pi(q_1, q_2, \cdots, q_n; \epsilon)$ defined on a *lattice* Q is supermodular if and only if for all $i \neq j$, $\frac{\partial^2 \Pi}{\partial q_i \partial q_j} > 0$; furthermore, if $\frac{\partial^2 \Pi}{\partial q_i \partial \epsilon} > 0, \forall i$, then function Π has strictly increasing differences in (q, ϵ). Let $q(\epsilon) = \underset{q \in Q}{max} \Pi(q, \epsilon)$. Then for a supermodular function with increasing differences in (q, ϵ), $q_i(\epsilon)$ is a strictly increasing function of ϵ for all i.

[5] Otherwise, the entrants can share the present network with the incumbent monopolist.

We assume that the network is congestible : $\Psi''(\cdot) < 0$. Besides, all the regular conditions $V'(\cdot) > 0, V''(\cdot) < 0$ and negative definiteness of Hessian matrix Π_{qq} in Proposition 1 are still satisfied.

The entry threat gives the consumers non-zero type-dependent reservation utilities, and thus the problem of the incumbent network supplier can be represented as

$$(\text{P5})\begin{cases} \max_{\{(\underline{U},\underline{q});(\overline{U},\overline{q})\}} v\left[\underline{\theta}V(\underline{q}) - c\underline{q}\right] + (1-v)\left[\overline{\theta}V(\overline{q}) - c\overline{q}\right] + \Psi(Q) - \left[v\underline{U} + (1-v)\overline{U}\right] \\ s.t.\ IR(\underline{\theta}) : \underline{U} \geqslant \underline{G} \\ \quad IR(\overline{\theta}) : \overline{U} \geqslant \overline{G} \\ \quad IC(\underline{\theta}) : \underline{U} \geqslant \overline{U} - \Delta\theta V(\overline{q}) \\ \quad IC(\overline{\theta}) : \overline{U} \geqslant \underline{U} + \Delta\theta V(\underline{q}). \end{cases}$$

Note that (P5) is the same as (P3) except for the non-zero type-dependent reservation utilities $\underline{G}$ and $\overline{G}$. The following proposition characterizes the optimal entry deterrence contract.

Proposition 4 *The optimal entry-deterring contract depends on the marginal cost of potential entrant. Specifically, there exist positive values $\omega_1 < \omega_2 < \omega_3 < \omega_4$ such that,*

1. *If $\omega > \omega_4$, then $\Delta G < \Delta\theta V(\underline{q}^{SB})$, and consequently the pricing contract is: $\underline{q} = \underline{q}^{SB}$, $\overline{q} = \overline{q}^{SB}$, $\underline{U} = \underline{G}$, and $\overline{U} = \underline{G} + \Delta\theta V(\underline{q}^{SB})$.*

2. *If $\omega_3 \leqslant \omega \leqslant \omega_4$, then $\Delta\theta V(\underline{q}^{SB}) \leqslant \Delta G \leqslant \Delta\theta V(\underline{q}^{FB})$, and consequently the optimal consumption level $\underline{q}$ and $\overline{q}$ are determined by*

$$\begin{cases} \underline{q} = V^{-1}\left(\dfrac{\Delta G}{\Delta\theta}\right) \\ \overline{\theta}V'(\overline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c, \end{cases} \tag{12}$$

   where $\underline{q} \in \left[\underline{q}^{SB}, \underline{q}^{FB}\right]$ and $\overline{q} \in [\overline{q}^{FB}, \overline{q}^{SB}]$. The consumers' information rents are $\underline{U} = \underline{G}$ and $\overline{U} = \overline{G}$.

3. *If $\omega_2 < \omega < \omega_3$, then $\Delta\theta V(\underline{q}^{FB}) < \Delta G < \Delta\theta V(\overline{q}^{FB})$, and consequently the pricing contract is $\underline{q} = \underline{q}^{FB}$, $\overline{q} = \overline{q}^{FB}$, $\underline{U} = \underline{G}$, and $\overline{U} = \underline{G}$.*

4. *If $\omega_1 \leqslant \omega \leqslant \omega_2$, then $\Delta\theta V(\overline{q}^{FB}) \leqslant \Delta G \leqslant \Delta\theta V(\overline{q}^{CI})$, and consequently the optimal consumption level $\underline{q}$ and $\overline{q}$ are given by*

$$\begin{cases} \overline{q} = V^{-1}\left(\dfrac{\Delta G}{\Delta\theta}\right) \\ \underline{\theta}V'(\underline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c, \end{cases} \tag{13}$$

where $\underline{q} \in [\underline{q}^{CI}, \underline{q}^{FB}]$ *and* $\overline{q} \in [\overline{q}^{FB}, \overline{q}^{CI}]$.[6] *The consumers' information rents are* $\underline{U} = \underline{G}$ *and* $\overline{U} = \overline{G}$.

5. *If* $0 < \omega < \omega_1$*, then* $\Delta G > \Delta\theta V(\overline{q}^{CI})$*, and consequently the optimal contract is* $\underline{q} = \underline{q}^{CI}$*,* $\overline{q} = \overline{q}^{CI}$*,* $\underline{U} = \overline{G} - \Delta\theta V(\overline{q}^{CI})$*,* $\overline{U} = \overline{G}$*.* $\underline{q}^{CI}$ *and* $\overline{q}^{CI}$ *are given by:*

$$\begin{cases} \underline{\theta}V'(\underline{q}^{CI}) + \Psi'\left(v\underline{q}^{CI} + (1-v)\overline{q}^{CI}\right) = c \\ \left(\overline{\theta} + \dfrac{v}{1-v}\Delta\theta\right)V'(\overline{q}^{CI}) + \Psi'\left(v\underline{q}^{CI} + (1-v)\overline{q}^{CI}\right) = c. \end{cases} \tag{14}$$

Proof. See Appendix ■

Remark 2 When $\omega > \omega_4$, the second best contract is also entry deterring. It means that when the outside competitors are not efficient enough to give high demand consumers enough utility exceeding their information rents acquired from the present network, the outside market is only attractive to low demand consumers. The incumbent firm need not to change its pricing contract when facing the entry threat of a firm with low competitiveness.

When $\omega_3 \leqslant \omega \leqslant \omega_4$, we have $\underline{q} \in \left[\underline{q}^{SB}, \underline{q}^{FB}\right]$ and $\overline{q} \in \left[\overline{q}^{FB}, \overline{q}^{SB}\right]$. That means when the marginal cost ω decreases to the extent that the utility difference ΔG is large enough to attract high demand consumers bypassing the present network, the monopolist must give up more information rent to him by increasing the consumption level of low demand consumers. The consumption level of high demand consumers themselves should also be lowered accordingly because of network effects. In this case, the sharper competitiveness of outside competitors makes the allocations less distorted in a sense that they are closer to the first best outcome than the second best outcome.

When $\omega_2 < \omega < \omega_3$, asymmetric information imposes no distortion on both types' allocation. As ω decreases and ΔG increases further, $\underline{q}$ will reach the first best level, it is suboptimal for the monopolist to increase the high type consumers' information rent at the cost of distorting the consumption level of the low demand consumers upward. In this case, the main task for the firm toward the high type consumers is to prevent them from bypassing the incumbent market instead of preventing them from misreporting, the participation constraints are more difficult to be satisfied than the incentive compatible constraints. Thus only the IRs are binding, and the first best allocation is attained.

When $\omega_1 \leqslant \omega \leqslant \omega_2$, we have $\underline{q} \in [\underline{q}^{CI}, \underline{q}^{FB}]$ and $\overline{q} \in [\overline{q}^{FB}, \overline{q}^{CI}]$. The high difference of utilities induces the low type consumers to pretend to be a high type, from which the countervailing incentives problem arises. The $IC(\underline{\theta})$, $IR(\underline{\theta})$ and $IR(\underline{\theta})$ in (P5) are binding. Again, the

[6] " CI " denotes " countervailing incentives ".

allocations of the two types will be distorted in opposite directions. But it is different from the distortions in cases 1 and 2. In these cases, the monopolist distorts $\underline{q}$ downward to curb the rent of high demand consumers. The upward distortion of $\overline{q}$ is for the purpose of counteracting the downward distortion of $\underline{q}$. In this regime, however, information rent has to be given to low demand consumers to elicit them reporting their types truthfully. The information rent is a decreasing function of high demand consumers' consumption $\overline{q}$, and so $\overline{q}$ has to be distorted upward to reduce the information rent gained by low demand consumers.

When $0 < \omega < \omega_1$, the allocations remain at the countervailing incentives level: $\underline{q} = \underline{q}^{CI}$ and $\overline{q} = \overline{q}^{CI}$. The decrease in marginal cost ω demand further upward distortion on the consumption of high demand consumers (the consumption of low demand consumers will be distorted downward accordingly). Thus, the participation constraint of the low-type has to be slackened, which means certain amount of information rent should be given to the consumers with low willingness to pay. In this case, only $IC(\underline{\theta})$ and $IR(\overline{\theta})$ are binding constraints, the low-type consumers get information rent $\overline{G} - \Delta\theta V\left(\overline{q}^{CI}\right)$. The incumbent firm keep reducing the tariffs ($\underline{t}$ and $\overline{t}$ keep decreasing in this case) instead of distorting allocations to prevent low demand consumers from bypassing and high demand consumers from misreporting.

5 Conclusion

This paper presents a model of nonlinear pricing in the presence of network externalities, under asymmetric information, and with the threat of entry. We give a full characterization of the nonlinear pricing contract. We showed that in un-congestible network, network externalities and asymmetric information together lead to a downward distortion for all consumers' consumption level; while in a congestible network, the consumptions of different types of consumers will be distorted in opposite directions. It can be regarded as a generalization of Hahn (2003), Segal (1999, 2003) and Csobra (2008).

We also analyzed the countervailing incentives issue resulted from the monopolist's entry deterring strategy. Entry deterrence appears to play a significant role in practice. Although the potential entrants have not yet formed their own consumers network, their low production cost can make them rivals to the incumbent monopolist. We discuss in detail how the entrants' marginal cost affects the pricing strategy of the incumbent. When the marginal cost is relative high, the pricing strategy is unaffected. As their competitiveness becoming sharper, the consumptions are less distorted, and there is no distortion ($\omega \in (\omega_1, \omega_2)$) in some interval for both types. When the marginal cost is low enough, the consumptions remain at the *countervailing* *incentives* level. Thus, the interdependency of both types is a notable feature which is different from the classical network bypassing literature.

Appendix

Proof of Lemma:

First note that the participation constraint IR_i will be automatically satisfied for all types $i \geqslant 2$ if the constraints IR_1 and IC_{i1} are satisfied since

$$U_i \geqslant U_1 + (\theta_i - \theta_1)V(q_1) > U_1 \geqslant 0.$$

Adding the constraints

$$IC_{ij} : U_i \geqslant U_j + (\theta_i - \theta_j)V(q_j)$$

and

$$IC_{ji} : U_j \geqslant U_i + (\theta_j - \theta_i)V(q_i),$$

we have

$$(\theta_i - \theta_j)[V(q_i) - V(q_j)] \geqslant 0.$$

Thus, in order to get an implementable mechanism, the consumption $q(\theta)$ should be a non-decreasing function in type. If $i > j$, the downward constraint IC_{ij} and the implementability condition $q_i > q_j$ imply the upward constraint IC_{ji}, but not vice versa. Thus only the downward ICs are binding. Adding

$$IC_{i(i-1)} : U_i \geqslant U_{i-1} + \Delta\theta V(q_{i-1})$$

and

$$IC_{(i-1)(i-2)} : U_{i-1} \geqslant U_{i-2} + \Delta\theta V(q_{i-2}),$$

we have

$$U_i \geqslant U_{i-2} + \Delta\theta \left[V(q_{i-1}) + V(q_{i-2})\right] \geqslant U_{i-2} + (\theta_i - \theta_{i-2})\, V(q_{i-2}),$$

in which the second inequality is ensured by the implementability condition. So the local incentive compatible constraints imply the non-local ones.

Proof of proposition 1:

The first order condition to (5) is:

$$\Pi_q(q, \alpha) = 0, \tag{15}$$

that implies,

$$\begin{cases} \alpha V'(\underline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c, \\ \overline{\theta} V'(\overline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c. \end{cases} \tag{16}$$

Differentiating (15) with respect to parameter α, we get:

$$\Pi_{qq}\frac{\mathrm{d}q}{\mathrm{d}\alpha}+\Pi_{q\alpha}=0 \tag{17}$$

that is,

$$\begin{pmatrix} \alpha v V''(\underline{q})+v^2\Psi''(Q) & v(1-v)\Psi''(Q) \\ v(1-v)\Psi''(Q) & (1-v)\overline{\theta}V''(\overline{q})+(1-v)^2\Psi''(Q) \end{pmatrix}\begin{pmatrix} \frac{\mathrm{d}\underline{q}}{\mathrm{d}\alpha} \\ \frac{\mathrm{d}\overline{q}}{\mathrm{d}\alpha} \end{pmatrix}+\begin{pmatrix} vV'(\underline{q}) \\ 0 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Solving the above equations, we have

$$\begin{cases} \dfrac{\mathrm{d}\underline{q}}{\mathrm{d}\alpha}=\dfrac{-V'(\underline{q})\left[\overline{\theta}V''(\overline{q})+(1-v)\Psi''(Q)\right]}{\alpha V''(\underline{q})\left[\overline{\theta}V''(\overline{q})+(1-v)\Psi''(Q)\right]+v\overline{\theta}V''(\overline{q})\Psi''(Q)} \\ \dfrac{\mathrm{d}\overline{q}}{\mathrm{d}\alpha}=\dfrac{vV'(\underline{q})\Psi''(Q)}{\alpha V''(\underline{q})\left[\overline{\theta}V''(\overline{q})+(1-v)\Psi''(Q)\right]+v\overline{\theta}V''(\overline{q})\Psi''(Q)} \\ \dfrac{\mathrm{d}Q}{\mathrm{d}\alpha}=v\dfrac{\mathrm{d}\underline{q}}{\mathrm{d}\alpha}+(1-v)\dfrac{\mathrm{d}\overline{q}}{\mathrm{d}\alpha}=\dfrac{-v\overline{\theta}V'(\underline{q})V''(\overline{q})}{\alpha V''(\underline{q})\Big[\overline{\theta}V''(\overline{q})+(1-v)\Psi''(Q)\Big]+v\overline{\theta}V''(\overline{q})\Psi''(Q)} \end{cases}. \tag{18}$$

From the fact that the Hessian matrix Π_{qq} is negative definite, it can be verified that the $2th$ diagonal element of Π_{qq} is negative, and thus

$$\overline{\theta}V''(\overline{q})+(1-v)\Psi''(Q)<0 \tag{19}$$

and the determinant of Π_{qq} is positive,

$$\det(\Pi_{qq})=v(1-v)\Big\{\alpha V''(\underline{q})\left[\overline{\theta}V''(\overline{q})+(1-v)\Psi''(Q)\right]+v\overline{\theta}V''(\overline{q})\Psi''(Q)\Big\}>0. \tag{20}$$

The signs of derivatives in (18) can be determined, which are $\frac{\mathrm{d}\underline{q}}{\mathrm{d}\alpha}>0$ and $\frac{\mathrm{d}Q}{\mathrm{d}\alpha}>0$. That means $\underline{q}^{SB}<\underline{q}^{FB}$ and $Q^{SB}<Q^{FB}$. The sign of $\frac{\mathrm{d}\overline{q}}{\mathrm{d}\alpha}$ depends on the sign of $\Psi''(Q)$: if $\Psi''(Q)>0$, $\frac{\mathrm{d}\overline{q}}{\mathrm{d}\alpha}>0$, then $\overline{q}^{SB}<\overline{q}^{FB}$; if $\Psi''(Q)<0$, $\frac{\mathrm{d}\overline{q}}{\mathrm{d}\alpha}>0$, then $\overline{q}^{SB}>\overline{q}^{FB}$.

Proof of proposition 2:

Expression (9) can be attained directly from the first order condition of (8). The weakly hazard rate property $H'(\theta)\leqslant 0$ ensures the implementability conditions $q_i^{SB}\leqslant q_{i+1}^{SB}, \forall i\in\{1,2,\cdots,n\}$, and the negative definiteness of Hessian matrix Π_{qq} ensures the second order sufficient condition of optimization. Thus (9) gives the second-best consumptions. The second best information rents and tariffs can be attained accordingly.

Proof of proposition 3:

Let

$$\Gamma \equiv \begin{pmatrix} [\theta_1+\epsilon H(\theta_1)]\, f(\theta_1)V''(q_1) & & & \\ & [\theta_2+\epsilon H(\theta_2)]\, f(\theta_2)V''(q_2) & & \\ & & \ddots & \\ & & & [\theta_n+\epsilon H(\theta_n)]\, f(\theta_n)V''(q_n) \end{pmatrix}$$

$$\gamma \equiv \big(f(\theta_1), f(\theta_2), ..., f(\theta_n)\big)^T.$$

Then

$$\Pi_{qq} \equiv \Gamma + \Psi''(Q)\gamma\gamma^T \tag{21}$$

$$\Pi_{q\epsilon} \equiv \Big(H(\theta_1)f(\theta_1)V'(q_1), H(\theta_2)f(\theta_2)V'(q_2), ..., H(\theta_n)f(\theta_n)V'(q_n)\Big)^T \tag{22}$$

The fist-order condition to (8) is $\Pi_q = 0$. The negative definiteness of Hessian matrix Π_{qq} acts as the second-order sufficient condition of the principal's optimization program, and the monotone hazard rate property ensures the implementable conditions $q_{i+1} \geqslant q_i, \forall i$.

Differentiating the above first-order condition with respect to parameter ϵ, we attain

$$\Pi_{qq}\frac{\mathrm{d}q}{\mathrm{d}\epsilon} + \Pi_{q\epsilon} = 0, \tag{23}$$

which implies

$$\frac{\mathrm{d}q}{\mathrm{d}\epsilon} = -(\Pi_{qq})^{-1}\Pi_{q\epsilon}. \tag{24}$$

Substituting expressions (21) and (22) into the above expression, we have

$$\frac{\mathrm{d}q}{\mathrm{d}\epsilon} = -\Big[\Gamma + \Psi''(Q)\gamma\gamma'\Big]^{-1}\Pi_{q\epsilon} = -\left[\Gamma^{-1} - \Psi''(Q)\frac{\Gamma^{-1}\gamma\cdot\gamma'\Gamma^{-1}}{1+\Psi''(Q)\gamma'\Gamma^{-1}\gamma}\right]\Pi_{q\epsilon}. \tag{25}$$

The *ith* elements of the LHS is then given by

$$\begin{aligned} \frac{\mathrm{d}q_i}{\mathrm{d}\epsilon} =& -\frac{H(\theta_i)V'(q_i)}{[\theta_i+\epsilon H(\theta_i)]V''(q_i)} \\ &+ \frac{1}{[\theta_i+\epsilon H(\theta_i)]V''(q_i)}\left\{\frac{\Psi''(Q)\sum_{j=1}^n \frac{H(\theta_j)f(\theta_j)V'(q_j)}{[\theta_j+\epsilon H(\theta_j)]V''(q_j)}}{1+\Psi''(Q)\sum_{j=1}^n \frac{f(\theta_j)}{[\theta_j+\epsilon H(\theta_j)]V''(q_j)}}\right\} \\ =& \frac{\rho - H(\theta_i)V'(q_i)}{[\theta_i+\epsilon H(\theta_i)]\,V''(q_i)}, \end{aligned} \tag{26}$$

where

$$\rho \equiv \frac{\Psi''(Q)\sum_{j=1}^n \frac{f(\theta_j)H(\theta_j)V'(q_j)}{[\theta_j+\epsilon H(\theta_j)]V''(q_j)}}{1+\Psi''(Q)\sum_{j=1}^n \frac{f(\theta_j)}{[\theta_j+\epsilon H(\theta_j)]V''(q_j)}}.$$

Note that the Hessian matrix Π_{qq} and hence its inverse Π_{qq}^{-1} are negative definite. So for non-zero vector γ we have

$$\begin{aligned}\gamma^T \Pi_{qq}^{-1} \gamma &= \gamma^T \left[\Gamma^{-1} - \Psi''(Q) \frac{\Gamma^{-1}\gamma \cdot \gamma^T \Gamma^{-1}}{1 + \Psi''(Q)\gamma^T \Gamma^{-1}\gamma} \right] \gamma \\ &= \frac{\gamma^T \Gamma^{-1}\gamma}{1 + \Psi''(Q)\gamma^T \Gamma^{-1}\gamma} < 0.\end{aligned}$$

$V''(\cdot) < 0$ implies every element of the diagonal matrix Γ^{-1} is negative, so $\gamma^T \Gamma^{-1} \gamma < 0$, and thus the denominator of ρ is positive, namely,

$$1 + \Psi''(Q)\gamma^T \Gamma^{-1}\gamma = 1 + \Psi''(Q) \sum_{j=1}^{n} \frac{f(\theta_j)}{[\theta_j + \alpha H(\theta_j)]V''(q_j)} > 0.$$

The sign of ρ and hence the sign of $\frac{\mathrm{d}q_i}{\mathrm{d}\epsilon}$ are determined by the sign of $\Psi''(Q)$.

1. If $\Psi''(Q) > 0$, then $\rho < 0$, $\frac{\mathrm{d}q_i}{\mathrm{d}\epsilon} > 0$, which implies $q_i^{SB} < q_i^{FB}$.

2. If $\Psi''(Q) < 0$, then $\rho > 0$. From conditions $H'(\theta) \leqslant 0$, $q_i \leqslant q_{i+1}$, and $V''(\cdot) < 0$, we get $H(\theta_1)V'(q_1) > H(\theta_2)V'(q_2) > \cdots > H(\theta_n)V'(q_n)$, then we have the following inequalities:

$$0 = H(\theta_n)V'(q_n) < \rho < \frac{\sum_{j=1}^{n} \frac{f(\theta_j)H(\theta_j)V'(q_j)}{[\theta_j + \alpha H(\theta_j)]V''(q_j)}}{\sum_{j=1}^{n} \frac{f(\theta_j)}{[\theta_j + \alpha H(\theta_j)]V''(q_j)}} < H(\theta_1)V'(q_1).$$

   Then there exists a unique $i^* \in \{1, 2, \cdots, n\}$ such that, when $i > i^*$, we have $\rho > H(\theta_i)V'(q_i)$, $\frac{\mathrm{d}q_i}{\mathrm{d}\epsilon} < 0$, and consequently $q_i^{SB} > q_i^{FB}$; when $i < i^*$, we have $\rho < H(\theta_i)V'(q_i)$, $\frac{\mathrm{d}q_i}{\mathrm{d}\epsilon} > 0$, and consequently $q_i^{SB} < q_i^{FB}$.

The derivative of Q with respective to ϵ is

$$\begin{aligned}\frac{\mathrm{d}Q}{\mathrm{d}\epsilon} &= \sum_{i=1}^{n} \left\{ \frac{[\rho - H(\theta_i)V'(q_i)]\, f(\theta_i)}{[\theta_i + \epsilon H(\theta_i)]\, V''(q_i)} \right\} \\ &= \frac{\Psi''(Q) \sum_{j=1}^{n} \frac{f(\theta_j)H(\theta_j)V'(q_j)}{[\theta_j + \epsilon H(\theta_j)]V''(q_j)}}{1 + \Psi''(Q) \sum_{j=1}^{n} \frac{f(\theta_j)}{[\theta_j + \epsilon H(\theta_j)]V''(q_j)}} \sum_{i=1}^{n} \frac{f(\theta_i)}{[\theta_i + \epsilon H(\theta_i)]\, V''(q_i)} - \sum_{i=1}^{n} \frac{f(\theta_i)H(\theta_i)V'(q_i)}{[\theta_i + \epsilon H(\theta_i)]\, V''(q_i)} \\ &= -\frac{\sum_{i=1}^{n} \frac{f(\theta_i)H(\theta_i)V'(q_i)}{[\theta_i + \epsilon H(\theta_i)]V''(q_i)}}{1 + \Psi''(Q) \sum_{i=1}^{n} \frac{f(\theta_i)}{[\theta_i + \epsilon H(\theta_i)]V''(q_i)}} > 0.\end{aligned} \tag{27}$$

Thus, we have $Q^{SB} < Q^{FB}$ for both cases.

Proof of Proposition 4:

In (P5), we have as many regimes as combinations of binding constraints among $IR(\underline{\theta})$, $IR(\overline{\theta})$, $IC(\underline{\theta})$ and $IC(\overline{\theta})$. To reduce the number of possible cases, we first give the following lemmas.

Lemma 2 *A pooling contract with $\underline{q} = \overline{q}$ and $\underline{t} = \overline{t}$ can never be optimal.*

Proof. Suppose that the optimal contract is pooling with $\underline{q} = \overline{q} = q$ and $\underline{t} = \overline{t} = t$. There are two cases to be considered.

(i) $\overline{\theta}V'(q) > c$. Then, increase $\overline{q}$ by ε and the transfer by $\overline{\theta}V'(q)\varepsilon$, the $\overline{\theta}-$type consumers can remain indifferent. Since at (q,t) the marginal rate of substitution between q and t is higher for $\overline{\theta}-$type consumers, this new allocation is incentive compatible. This raises the firm's revenue by $(1-v)[\overline{\theta}V'(q) - c]\varepsilon$.

(ii) $\overline{\theta}V'(q) \leqslant c$ and $\underline{\theta}V'(q) < c$. Then, decrease $\underline{q}$ by ε and adjust $\underline{t}$ so that $\underline{\theta}-$type consumers can remain on the same indifference curve. Then the firm's total charge will be increased by $[c - \underline{\theta}V'(q)]\varepsilon$.

Thus, in both cases, it contradicts with the fact (q,t) is optimal contract. ■

Lemma 3 *If the two types of consumers are offered two different contracts, the two incentive constraints cannot be simultaneously bindings.*

Proof. Suppose by way of contradiction that both ICs are binding. From $\underline{\theta}V(\underline{q}) - \underline{t} + \Psi(Q) = \underline{\theta}V(\overline{q}) - \overline{t} + \Psi(Q)$ and $\overline{\theta}V(\overline{q}) - \overline{t} + \Psi(Q) = \overline{\theta}V(\underline{q}) - \underline{t} + \Psi(Q)$, we have $\underline{q} = \overline{q}$ and $\underline{t} = \overline{t}$. But this is impossible by Lemma 1. ■

Lemma 4 *The IC and IR constraints of the same type cannot be simultaneously slack.*

Proof. If $IR(\theta)$ and $IC(\theta)$ are both slack, increase $t(\theta)$ by a tiny increment will not violate all the constraints, but the firm's charge will be increased. ■

Applying the above three lemmas, only five possible regimes are needed to be considered, which are summarized in the following table.

Table1 − The five possible regimes

Constraints	Regime1	Regime2	Regime3	Regime4	Regime5
$IR(\underline{\theta})$	B	B	B	B	S
$IR(\overline{\theta})$	S	B	B	B	B
$IC(\underline{\theta})$	S	S	S	B	B
$IC(\overline{\theta})$	B	B	S	S	S

Where "B" denotes "binding" and "S" denotes "slack".

The regimes are ordered from 1 to 5 when ΔG increases, and ΔG itself is determined by the entrants' marginal cost ω. To find how ω affects the nonlinear pricing contract of the incumbent firm, we give the following two lemmas. Lemma 5 states the the variation of ΔG in different regimes, and Lemma 6 shows how ΔG is determined by ω.

Lemma 5 *The optimal pricing contracts and the utility difference ΔG in different regimes are:*

1. *In regime 1, the optimal solution to (P5) is $\underline{q} = \underline{q}^{SB}$, $\overline{q} = \overline{q}^{SB}$, $\underline{U} = \underline{G}$, and $\overline{U} = \underline{G} + \Delta\theta V(\underline{q}^{SB})$. The value of utility difference satisfies $\Delta G < \Delta\theta V(\underline{q}^{SB})$.*

2. *In regime 2, the optimal consumption level $\underline{q}$ and $\overline{q}$ are determined by*

$$\begin{cases} \underline{q} = V^{-1}\left(\dfrac{\Delta G}{\Delta\theta}\right) \\ \overline{\theta}V'(\overline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c, \end{cases} \tag{28}$$

where $\underline{q} \in \left[\underline{q}^{SB}, \underline{q}^{FB}\right]$ and $\overline{q} \in [\overline{q}^{FB}, \overline{q}^{SB}]$. The consumers' information rents are $\underline{U} = \underline{G}$ and $\overline{U} = \overline{G}$. The utility difference satisfies $\Delta\theta V(\underline{q}^{SB}) \leqslant \Delta G \leqslant \Delta\theta V(\underline{q}^{FB})$.

3. *In regime 3, the optimal solution to (P5) is $\underline{q} = \underline{q}^{FB}$, $\overline{q} = \overline{q}^{FB}$, $\underline{U} = \underline{G}$, $\overline{U} = \overline{G}$, and $\Delta\theta V(\underline{q}^{FB}) < \Delta G < \theta V(\overline{q}^{FB})$.*

4. *In regime 4, the optimal consumption level $\underline{q}$ and $\overline{q}$ are determined by*

$$\begin{cases} \overline{q} = V^{-1}\left(\dfrac{\Delta G}{\Delta\theta}\right) \\ \underline{\theta}V'(\underline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c, \end{cases} \tag{29}$$

where $\underline{q} \in [\underline{q}^{CI},\ \underline{q}^{FB}]$ and $\overline{q} \in [\overline{q}^{FB}, \overline{q}^{CI}]$. The consumers' information rents are $\underline{U} = \underline{G}$ and $\overline{U} = \overline{G}$. The utility difference satisfies $\Delta\theta V(\overline{q}^{FB}) \leqslant \Delta G \leqslant \Delta\theta V(\overline{q}^{CI})$.

5. *In regime 5, the optimal contract is $\underline{q} = \underline{q}^{CI}$, $\overline{q} = \overline{q}^{CI}$, $\underline{U} = \overline{G} - \Delta\theta V(\overline{q}^{CI})$, and $\overline{U} = \overline{G}$. The utility difference $\Delta G > \theta V(\overline{q}^{CI})$. $\underline{q}^{CI}$ and $\overline{q}^{CI}$ are determined by:*

$$\begin{cases} \underline{\theta}V'(\underline{q}^{CI}) + \Psi'\left(v\underline{q}^{CI} + (1-v)\overline{q}^{CI}\right) = c \\ \left(\overline{\theta} + \dfrac{v}{1-v}\Delta\theta\right)V'(\overline{q}^{CI}) + \Psi'\left(v\underline{q}^{CI} + (1-v)\overline{q}^{CI}\right) = c. \end{cases} \tag{30}$$

Proof.

- In regime 1, the constraints $IR(\underline{\theta})$ and $IC(\overline{\theta})$ are binding. Solving (P5) in the same way as (P3), we get the second best solution

$$\left\{\underline{q} = \underline{q}^{SB}, \overline{q} = \overline{q}^{SB}; \underline{U} = \underline{G}, \overline{U} = \underline{G} + \Delta\theta V(\underline{q}^{SB})\right\},$$

with $\Delta G < \Delta U = \Delta\theta V(\underline{q}^{SB})$.

- In regime 2, the constraints $IR(\underline{\theta})$, $IR(\overline{\theta})$ and $IC(\overline{\theta})$ are binding. The optimal contract set is thus given by

$$\Big\{(\underline{q},\overline{q},\underline{U},\overline{U}) : \Delta\theta V(\underline{q}) = \Delta G, \overline{\theta}V(\overline{q}) + \Psi'(Q) = c; \underline{U} = \underline{G}, \overline{U} = \overline{G}\ \Big\}.$$

Substituting $IR(\underline{\theta})$ and $IR(\overline{\theta})$ into the objective function , the Lagrange function of is constructed as

$$\begin{aligned} L(\underline{q},\overline{q}) &= v\Big[\underline{\theta}V(\underline{q}) - c\underline{q}\Big] + (1-v)\Big[\overline{\theta}V(\overline{q}) - c\overline{q}\Big] \\ &\quad + \Psi(Q) - \Big[v\underline{G} + (1-v)\overline{G}\Big] + \lambda\Big[\Delta G - \Delta\theta V(\underline{q})\Big], \end{aligned}$$

where $\lambda > 0$ is the Lagrange multiplier of the binding constraint $IC(\overline{\theta})$. Then $\underline{q}$ and $\overline{q}$ are determined by:

$$\begin{cases} \left(\underline{\theta} - \dfrac{\lambda}{v}\Delta\theta\right) V'(\underline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c, \\ \overline{\theta}V'(\overline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c. \end{cases} \tag{31}$$

Because $\underline{\theta} - \frac{\lambda}{v}\Delta\theta < \underline{\theta}$, from formula (18) it is easy to verify that $\underline{q} < \underline{q}^{FB}$ and $\overline{q} > \overline{q}^{FB}$. Substituting $IR(\underline{\theta})$ and $IC(\overline{\theta})$ into the objective function of (P5) and letting $\delta > 0$ be the Lagrange multiplier associate with the binding constraint $IR(\overline{\theta})$, we obtain the following Lagrange function:

$$\begin{aligned} L(\underline{q},\overline{q}) &= v\Big[\underline{\theta}V(\underline{q}) - c\underline{q}\Big] + (1-v)\Big[\overline{\theta}V(\overline{q}) - c\overline{q}\Big] \\ &\quad + \Psi(Q) - \Big[v\underline{G} + (1-v)(\underline{G} + \Delta\theta V(\underline{q}))\Big] + \delta\Big[\Delta\theta V(\underline{q}) - \Delta G\Big]. \end{aligned}$$

The optimal consumptions $\underline{q}$ and $\overline{q}$ are determined by:

$$\begin{cases} \left(\underline{\theta} - \dfrac{1-v-\delta}{v}\Delta\theta\right) V'(\underline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c, \\ \overline{\theta}V'(\overline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c. \end{cases} \tag{32}$$

Because $\underline{\theta} - \frac{1-v-\delta}{v}\Delta\theta > \underline{\theta} - \frac{1-v}{v}\Delta\theta$, from expression (18) we have $\underline{q} > \underline{q}^{SB}$ and $\overline{q} < \overline{q}^{SB}$. It then suffice to show that $\Delta\theta V(\underline{q}^{SB}) \leqslant \Delta G \leqslant \Delta\theta V(\underline{q}^{FB})$.

- In regime 3, $IR(\underline{\theta})$ and $IR(\overline{\theta})$ are binding. Then the optimal contract is given by

$$\Big\{\underline{q} = \underline{q}^{FB}, \overline{q} = \overline{q}^{FB}; \underline{U} = \underline{G}, \overline{U} = \overline{G}\ \Big\}.$$

From the slack ICs, we can verify that ΔG satisfies $\Delta\theta V(\underline{q}^{FB}) < \Delta G < \Delta\theta V(\overline{q}^{FB})$.

- In regime 4, $IR(\underline{\theta})$ $IR(\overline{\theta})$ and $IC(\underline{\theta})$ are binding. The optimal contract is:

$$\Big\{(\underline{q},\overline{q},\underline{U},\overline{U}) : \Delta\theta V(\overline{q}) = \Delta G, \underline{\theta} V(\underline{q}) + \Psi'(Q) = c; \underline{U} = \underline{G}, \overline{U} = \overline{G}\ \Big\}.$$

Substituting $IR(\underline{\theta})$ and $IR(\overline{\theta})$ into the objective function, and letting $\mu > 0$ the multiplier associate with the binding constraint $IC(\underline{\theta})$, we have the following Lagrange function:

$$\begin{aligned} L(\underline{q},\overline{q}) = {} & v\Big[\underline{\theta}V(\underline{q}) - c\underline{q}\Big] + (1-v)\Big[\overline{\theta}V(\overline{q}) - c\overline{q}\Big] \\ & + \Psi(Q) - \Big[v\underline{G} + (1-v)\overline{G}\Big] + \mu\Big[\Delta\theta V(\overline{q}) - \Delta G\Big]. \end{aligned}$$

Thus, $\underline{q}$ and $\overline{q}$ are determined by

$$\begin{cases} \underline{\theta}V'(\underline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c, \\ \left(\overline{\theta} + \dfrac{\mu}{1-v}\Delta\theta\right) V'(\overline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c. \end{cases} \tag{33}$$

Substituting $IR(\overline{\theta})$ and $IC(\underline{\theta})$ into the objective function, and letting $\eta > 0$ be the multiplier associate with the binding constraint $IR(\underline{\theta})$, we have the Lagrange function:

$$\begin{aligned} L(\underline{q},\overline{q}) = {} & v\Big[\underline{\theta}V(\underline{q}) - c\underline{q}\Big] + (1-v)\Big[\overline{\theta}V(\overline{q}) - c\overline{q}\Big] + \Psi(Q) \\ & - \Big[v(\overline{G} - \Delta\theta V(\overline{q})) + (1-v)\overline{G}\Big] + \eta\Big[\Delta G - \Delta\theta V(\overline{q})\Big]. \end{aligned}$$

Then the $\underline{q}$ and $\overline{q}$ are determined by:

$$\begin{cases} \underline{\theta}V'(\underline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c, \\ \left(\overline{\theta} + \dfrac{v-\eta}{1-v}\Delta\theta\right) V'(\overline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c. \end{cases} \tag{34}$$

To compare the different consumption levels, we make some comparative static analysis. To do so, let

$$\begin{cases} \underline{\theta}V'(\underline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c, \\ \beta V'(\overline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c. \end{cases} \tag{35}$$

If $\beta = \overline{\theta}$, it's the expression of $\underline{q}^{FB}$ and $\overline{q}^{FB}$; if $\beta = \overline{\theta} + \frac{v}{1-v}\Delta\theta$, it coincides with the countervailing incentives solution $\underline{q}^{CI}$ and $\overline{q}^{CI}$.

Differentiating these two equations with respect to β leads to:

$$\begin{cases} \left[\underline{\theta}V''(\underline{q}) + v\Psi''(Q)\right]\dfrac{\mathrm{d}\underline{q}}{\mathrm{d}\beta} + (1-v)\Psi''(Q)\dfrac{\mathrm{d}\overline{q}}{\mathrm{d}\beta} = 0, \\ v\Psi''(Q)\dfrac{\mathrm{d}\underline{q}}{\mathrm{d}\beta} + \left[\beta V''(\overline{q}) + (1-v)\Psi''(Q)\right]\dfrac{\mathrm{d}\overline{q}}{\mathrm{d}\beta} = -V'(\underline{q}). \end{cases} \tag{36}$$

Thus, when $\Psi''(Q) < 0$ we have

$$\begin{cases} \dfrac{\mathrm{d}\underline{q}}{\mathrm{d}\beta} = \dfrac{(1-v)V'(\underline{q})\Psi''(Q)}{\beta V''(\underline{q})[\underline{\theta}V''(\underline{q}) + v\Psi''(Q)] + (1-v)\underline{\theta}V''(\underline{q})\Psi''(Q)} < 0, \\ \dfrac{\mathrm{d}\overline{q}}{\mathrm{d}\beta} = \dfrac{-V'(\underline{q})[\underline{\theta}V''(\underline{q}) + v\Psi''(Q)]}{\beta V''(\underline{q})[\underline{\theta}V''(\underline{q}) + v\Psi''(Q)] + (1-v)\underline{\theta}V''(\underline{q})\Psi''(Q)} > 0. \end{cases} \tag{37}$$

Because $\overline{\theta} + \frac{\mu}{1-v}\Delta\theta > \overline{\theta}$ and $\overline{\theta} + \frac{v-\eta}{1-v}\Delta\theta < \overline{\theta} + \frac{v}{1-v}\Delta\theta$, from formula (37), it can be verified that $\overline{q} > \overline{q}^{FB}$, $\underline{q} < \underline{q}^{FB}$, $\overline{q} < \overline{q}^{CI}$, and $\underline{q} > \underline{q}^{CI}$. Thus, $\Delta G = \Delta\theta V(\overline{q}) \in [\Delta\theta V(\overline{q}^{FB}), \Delta\theta V(\overline{q}^{CI})]$.

- In regime 5, $IR(\overline{\theta})$ and $IC(\underline{\theta})$ are binding constraints. Substituting $\overline{U} = \overline{G}$ and $\underline{U} = \overline{G} - \Delta\theta V(\overline{q})$ into the objective function, we obtain the following first order conditions:

$$\begin{cases} \underline{\theta}V'(\underline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c, \\ \left(\overline{\theta} + \dfrac{v}{1-v}\Delta\theta\right)V'(\overline{q}) + \Psi'\left(v\underline{q} + (1-v)\overline{q}\right) = c. \end{cases} \tag{38}$$

  It is the countervailing incentives consumption level. Note that $\overline{\theta} + \frac{v}{1-v}\Delta\theta > \overline{\theta}$, and thus, from (37) $\underline{q}^{FB} > \underline{q}^{CI}$, $\overline{q}^{FB} < \overline{q}^{CI}$. The difference of reservation utility satisfies $\Delta G > \Delta U = \Delta\theta V(\overline{q}^{CI})$.

Figures 1-5 give a graphical depiction of the above five regimes. ■

Lemma 6 *Suppose $V(0) = 0$, $V'(\cdot) > 0$, $V''(\cdot) < 0$, and $V(\cdot)$ satisfies the standard Inada conditions:* $\lim_{q\to+\infty}V'(q) = 0$, *and* $\lim_{q\to 0}V'(q) = +\infty$. *Then the utility difference across different states $\Delta G = \overline{G} - \underline{G}$ is a decreasing function of the marginal cost ω,* $\lim_{\omega\to 0}\Delta G = +\infty$, *and* $\lim_{\omega\to+\infty}\Delta G = 0$.

Proof. The first order condition of $G(\theta) = \max_q[\theta V(q) - \omega q]$ is given by $\theta V'(q^*) = \omega$. So the maximized utility derived from network bypassing is: $G(\theta) = \theta[V(q^*(\theta)) - q^*(\theta)V'(q^*(\theta))]$. Let $\Phi(q) = V(q) - qV'(q)$. Then $\Delta G = G(\overline{\theta}) - G(\underline{\theta}) = \overline{\theta}\Phi(\overline{q}^*) - \underline{\theta}\Phi(\underline{q}^*)$, and its derivative with respect to marginal cost ω is:

$$\begin{aligned} \frac{\mathrm{d}\Delta G}{\mathrm{d}\omega} &= \overline{\theta}\Phi'(\overline{q}^*)\frac{\mathrm{d}\overline{q}^*}{\mathrm{d}\omega} - \underline{\theta}\Phi'(\underline{q}^*)\frac{\mathrm{d}\underline{q}^*}{\mathrm{d}\omega} \\ &= -\overline{\theta}\overline{q}^*V''(\overline{q}^*)\frac{\mathrm{d}\overline{q}^*}{\mathrm{d}\omega} + \underline{\theta}\underline{q}^*V''(\underline{q}^*)\frac{\mathrm{d}\underline{q}^*}{\mathrm{d}\omega} \\ &= -\overline{\theta}\overline{q}^*V''(\overline{q}^*)\frac{1}{\overline{\theta}V''(\overline{q}^*)} + \underline{\theta}\underline{q}^*V''(\underline{q}^*)\frac{1}{\underline{\theta}V''(\underline{q}^*)} \\ &= -\overline{q}^* + \underline{q}^* < 0. \end{aligned}$$

It is easy to verify that when the conditions $V(0) = 0$, $V'(\cdot) \geqslant 0$, $V''(\cdot) < 0$, $\lim_{q\to+\infty}V'(q) = 0$, and $\lim_{q\to 0}V'(q) = +\infty$ are satisfied, $\lim_{\omega\to 0}\Delta G = +\infty$ and $\lim_{\omega\to+\infty}\Delta G = 0$. Figure 6 in appendix F depicts the relationship of ΔG and ω. ■

From the above lemmas, one can see that if the potential entrants' competitiveness increases, the utility differences will increase from zero to infinity. Thus, there exist positive values $\omega_i, i = 1, 2, 3, 4$, such that $\omega_1 < \omega_2 < \omega_3 < \omega_4$ corresponding to $\Delta\theta V(\overline{q}^{CI})$, $\Delta\theta V(\overline{q}^{FB})$, $\Delta\theta V(\underline{q}^{FB})$ and $\Delta\theta V(\underline{q}^{SB})$, respectively, where $\overline{q}^{CI}$, $\overline{q}^{FB}$, $\underline{q}^{FB}$ and $\underline{q}^{SB}$ are given in expressions (14), (3) and (4). Thus, combining Lemmas 5 and 6, we can easily prove the statement in Proposition 4. Figure 7 depicts the change of consumptions with ω.

References

[1] Baron, D., and R. Myerson. (1982), "Regulating a Monopolist with Unknown Costs", *Econometrica*, 50, 911-930.

[2] Csorba, G. (2008) : "Screening Contracts in The Presence of Positive Network Effects", *International Journal of Industrial Organization*, 26, 213-226 .

[3] Farrell, J. and Saloner, G. (1985), "Standardization, Compatibility, and Innovation" *Rand Journal of Economics*, 16, 70-83.

[4] Fudenberg, D. and Tirole, J.,(2000), "Pricing a Network Good to Deter Entry" *Journal of Industrial Economics*, 48, 373-390.

[5] Hahn, J-H.(2003), "Nonlinear Pricing of Telecommunication with Call and Network Externalities" *International Journal of Industrial Organization*, 21, 949-967.

[6] Katz, M. and Shapiro, C.,(1985), "Network Externalities, Competition and Contracting. " *American Economic Review*, 75, 424-440.

[7] Laffont. JJ, and Martimort. D, (2002) "The Theory Of Incentives–The Principal-Agent Model" *Princeton University Press*

[8] Laffont. JJ, and J Tirole, (1990), "Optimal Bypass and Cream Skimming" *The American Economic Review*, 80 ,1042-1061.

[9] Lewis, Tracy R. and Sappington, David E. M., (1989), "Countervailing Incentives in Agency Problems", *Journal of Economic Theory*, 49, 294-313

[10] Maggi,G. and A, Rodriguez-Clare., 1995"On countervailing incentives," *Journal of Economic Theory*, 66, 238-263.

[11] Maskin, E. and Riley, J.,(1984), "Monopoly with Incomplete Information", *Rand Journal of Economics*, 15, 171-196.

[12] Mendelson, H. (1985), "Pricing Computer Services: Queuing Effects" , *Communications of the ACM* ,28, 312-321.

[13] Mendelson, H. and Whang, S. (1990), "Optimal Incentive-Compatible Priority Pricing for the M/M/1 Queue" *Operations Research*, 38, 870-883.

[14] Milgrom, P. and Shannon, C. (1994), "Monotone Comparative Statistics", *Econometrica*, 62, 157-180

[15] Mussa, M. and Rosen, S. (1978), "Monopoly and Product Quality", *Journal of Economic Theory*, 18, 301-317

[16] Myerson, R. (1981), "Optimal Auction Design", *Mathematics of Operations Research*, 6, 58-73

[17] Oren, S., Smith, S., and Wilson, R. (1982), "Nonlinear Pricing in Markets with Interdependent Demand" , *Marketing Science*, 1, 287-313

[18] Rochet J-C, Stole L A., (2002), "Nonlinear pricing with random participation", *Review of Economic Studies*, 69, 277-311

[19] Rohlfs, J. (1974), "A Theory of Interdependent Demand for a Communication Service", *Bell Journal of Economics,* 10, 16-37.

[20] Segal, I., (1999), "Contracting with Externalities", *Quarterly Journal of Economics*, 114, 337-388

[21] Segal, I., (2003),, "Coordination and Discrimination in Contracting with Externalities: Divide and Conquer", *Journal of Economic Theory*, 113, 147-181

[22] Sundararajan. A., (2004), "Nonlinear Pricing and Type-dependent Network Effects" *Economics Letters,* 83, 107-113

[23] Topkis, D. (1998), "Supermodularity and Complementarity", *Princeton University Press.*

[24] Westland, J. Christopher.,(1992), "Congestion and Network Externalities in the Short Run Targeting of Information System Services, " *Management Science,* 38 , 992-1009.