Internet Interconnection and Network Neutrality

Jay Pil Choi[*] Doh-Shin Jeon[†] Byung-Cheol Kim[‡]

June 19, 2012

Abstract

This paper analyzes competition between interconnected networks when content is heterogeneous in terms of its sensitivity to delivery quality. In a two-sided market framework, we consider two broad regimes under which packet delivery can take place. Under a neutral regime mandated by net neutrality regulation, all packets are delivered with the same quality (speed). Under a non-neutral regime, Internet service providers (ISPs) are allowed to offer multiple lanes with different delivery quality levels. We show that the merit of net neutrality regulation depends crucially on content providers (CPs)' business models. The result also contributes more generally to the literature on second degree price discrimination by illustrating how second degree price discrimination fares against no discrimination depends on the nature of business models in a two-sided market.

JEL Classification: D4, L12, L4, L43, L51, L52

Key Words: Net neutrality, Internet interconnection, Two-sided markets, Second-degree price discrimination

[*]School of Economics, University of New South Wales, Sydney, NSW 2052, Australia and Department of Economics, Michigan State University, 220A Marshall-Adams Hall, East Lansing, MI 48824 -1038. E-mail: choijay@msu.edu.

[†]Toulouse School of Economics and CEPR., Manufacture de Tabacs, 21 allees de Brienne - 31000 Toulouse, France. E-mail: dohshin.jeon@gmail.com.

[‡]School of Economics, Georgia Institute of Technology, 221 Bobby Dodd Way, Atlanta, GA 30332-0225. Email: byung-cheol.kim@econ.gatech.edu.

1 Introduction

The Internet is a system of interconnected computer networks that is often characterized as a "network of networks." The universal connectivity that enables any computers connected to the Internet to communicate with each other is ensured by cooperative interconnection arrangements among network operators. The current state of the Internet is also governed by a more or less implicit principle of "net neutrality" that treats all packets equally and deliver them on a first-come-first-served basis without blocking or prioritizing any traffic based on types of Internet content, services or applications.

However, with the emergence of various on-line multi-media services that demand a significant amount of network bandwidth, network congestion and efficient management of network resources have become an important policy issue. In particular, content and applications differ in their sensitivity with respect to delay in delivery. For instance, data applications such as E-mail can be relatively insensitive toward moderate delivery delays from the users' viewpoints. By contrast, streaming video/audio or VoIP applications can be very sensitive to delay, leading to jittery delivery of content that provides unsatisfactory user experiences. With such heterogeneity concerning delay costs, one may argue that network neutrality treating all packets equally regardless of content is not an efficient way to utilize the network.

Even if there is an agreement concerning the desirability of offering multi-tiered Internet services, implementation of such a system is not a simple matter with interconnected networks. In particular, guaranteeing a specified quality (speed) of content delivery requires cooperation from other networks when content providers and end users belong to different networks. Though interconneted ISPs agree on the provision of delivery quality, they may well compete in the two groups of end users, consumers who subscribe the access to the Internet and content/application providers who want to deliver their content for businesses.

In this paper, we develop a theoretical model of interconnection to reflect these key features of the Internet ecosystem and highlight the importance of content providers' business models in assessing the effects of net neutrality. More specifically, we adopt a two-sided market framework in which ISPs serve as platforms that connect content providers (hereafter CPs) and end consumers. On the CP side, there is a continuum of heterogeneous content/application providers who can multi-home, i.e., subscribe to multiple ISPs. CPs' contents differ in their sensitivity to delivery quality. This justifies the need to provide multiple lanes of different delivery qualities. Consumers are assumed to single-home and constitute competitive bottlenecks in the market.[1] To model competition in the consumer side of the market, we employ a Hotelling model with hinterlands (Armstrong and Wright, 2009) to represent elastic subscription demand by consumers and ISPs' market power vis-a-vis consumers. On the contrary, we assume Bertrand competition without friction on the content side, which simplifies our analysis. These assumptions are made to reflect a typical real world environment in which ISPs have strong market power with respect to consumers because of the lack of competition for "the last mile" delivery while their market power is limited with respect to content providers who can choose among multiple ISPs to distribute their content.

When both CPs and consumers belong to the same ISP, all traffic can be delivered on-net. However, if a CP purchases a delivery service from one ISP and consumers subscribe to another ISP, interconnection between these two ISPs is required for the completion of content delivery. We consider two broad regimes under which packet delivery can take place. Under a neutral regime mandated by net neutrality regulation, all packets are delivered with the same quality (speed). Under a non-neutral regime, in contrast, ISPs are allowed to offer multiple lanes with different delivery quality levels. We assume that the ISPs agree on reciprocal access charges for the delivery of other ISPs' traffic that terminate on their own networks and the delivery quality. As in Laffont, Rey, and Tirole (1998a, 1998b), we further distinguish two cases in the non-neutral network, depending on whether ISPs can discriminate CPs based on the destination of content delivery. With termination-based price discrimination (TPD), ISPs can charge different prices to CPs for traffic that terminate on their own networks and traffic that terminate on rivals' networks. Without TPD, ISPs are required to charge the same price schedule regardless of the destination of the content delivery.

We find that any equilibrium in our model is governed by the so-called off-net cost pricing principle on the CP side, generalizing the results of Laffont et al. (2003) to a setting of heterogenous content with different delivery qualities across content. We establish that off-net cost pricing on the CP side combined with (ii)Hotelling competition with hinterland on the consumer side creates an equivalence between competing ISPs in our model and a hypothetical benchmark case of monopoly with homogenous consumers. Competing ISPs essentially agree on access charges and delivery qualities that would enable them to behave as monopoly bottlenecks with respect to CPs. By using this equivalence, we consider a scenario that would favor price discrimination and thus stack the deck against the neutral regime when the surpluses from interactions between the CPs and

[1] See Armstrong (2006).

end consumers are entirely appropriated by one-side of the market. Nonetheless, we show that a neutral regime can be welfare-enhancing when the degree of surplus extraction from the CPs is in the intermediate range. This result highlights the important of the CPs' business models in the evaluation of net neutrality regulation, which has not been considered in the debate. This result also contributes more generally to the literature on second degree price discrimination by illustrating how second degree price discrimination fares against no discrimination depends on the nature of business model in a two-sided market.

Our paper is closely related to Laffont, Marcus, Rey, and Tirole (hereafter LMRT, 2003) who analyze how the access charge allocates communication costs between CPs and end consumers and thus affects competitive strategies of rival networks in an environment of interconnected networks. They show that the principle of "off-net-cost pricing," in which network operators set prices for their customers as if their customers' traffic were entirely off-net, prevails in a broad set of environments. Our model builds upon their interconnection model, but focuses on the provision of optimal quality in content delivery services by introducing heterogeneity in CPs' content type. In addition, we analyze how the access charge is determined and how it impacts market competition in unregulated environments and compares the outcomes to the one under net neutrality regulation. There is a large literature on interconnection in the telecommunication market. Armstrong (1998) and Laffont, Rey, and Tirole (1998a), for instance, show that firms agree to set interconnection charges above associated costs in order to obtain the joint profit-maximizing outcome and derives the welfare-maximizing interconnection charge that is lower than the privately negotiated level. Laffont, Rey, and Tirole (1998 a, b) derive similar results for linear pricing, but they extend the analysis to the cases of two-part tariffs and termination-based price discrimination and demonstrate that the nature of competition can be altered significantly depending on whether two-part tariffs or price discrimination are employed or not as price instruments. Their models, however, are devoid of the issue of transmission quality because all calls are homogeneous. In contrast, our framework assumes heterogeneous types of CPs which require different transmission qualities in order to analyze the quality distortion associated with non-neutral networks and the (sub)optimality of net neutrality regulation.

Our paper also contributes to the literature on net neutrality. With net neutrality being one of the most important global regulatory issues concerning the Internet, there has been a steady stream of academic papers on various issues associated with net neutrality regulation in recent years. To the best of our knowledge, we are the first to explore implications of net neutrality in the framework of two-sided markets with *interconnected* and *competing* ISPs. Most papers on net neutrality consider a monopolistic ISP. Hermalin and Katz (2007) examine a situation in which ISPs serve as a platform to connect content providers with end consumers in a framework of two-sided markets. They consider heterogeneous content providers whose products are vertically differentiated in order to analyze the effects of net neutrality regulation. Without any restrictions, an ISP can potentially offer a continuum of vertically differentiated services, although the ISP is required to provide only one service (a single tier of Internet service) with net neutrality regulation. They compare the single service level equilibrium with the multi-service level equilibrium. Choi and Kim (2010) and Cheng, Bandyopadhyay and Guo (2009) analyze the effects of net neutrality regulation on investment incentives of ISPs and CPs. Economides and Hermalin (2012) derive conditions under which network neutrality would be welfare superior to any feasible scheme for prioritized service given a capacity of bandwidth. In their extension, they also show that the ability to price discriminate enhances incentives to invest, creating a trade-off between static and dynamic efficiencies. However, all these papers consider a monopolistic ISP and thus the issue of interconnection does not arise.[2]

In contrast, Bourreau, Kourandi, and Valletti (2012) analyze the effect of net neutrality regulation on capacity investments and innovation in the content market with competing ISPs. They show that investments in broadband capacity and content innovation are higher under a non-neutral regime. However, they do not allow interconnection between ISPs by assuming that a CP has access only to the end users connected to the same ISP. Economides and Tag (forthcoming) also consider both a monopolistic ISP and duopolistic ISPs. But once again, the issue of internet interconnection is not considered as they focus on how net neutrality regulation as zero pricing rule affects pricing schemes on both sides of the market and social welfare.

The remainder of the paper is organized in the following way. Section 2 sets up a basic model of interconnected networks with competition where the two ISPs play the role of a two-sided market platform that connects content providers on one side and end consumers on the other side. In section 3, we consider two benchmark cases of the first best and a monopolistic ISP with homogenous consumers against which the market equilibrium in various regimes can be compared. The latter case is intended to introduce some of the key parametric assumptions and establish

[2]Economides and Tag (forthcoming) also provide an economic analysis of net neutrality in a two-sided market framework and investigate how net neutrality regulation affects pricing schemes on both sides of the market and social welfare. They consider both a monopolistic ISP and duopolistic ISPs, but once again, the issue of internet interconnection is not considered.

connections between the monopoly case and competiting ISPs later. Sections 4 studies network competition in the CP side of the market and show that any equilibrium is characterized by the off-net cost pricing principle regardless of the regimes concerning neutrality and termination based price discrimination. Section 5 analyzes network competition in consumer subscription market and derives a central equivalnece result between competing ISPs and a monopolistic ISP: the ISPs that agree to access charges and delivery qualities to maximize their joint profits behave as a monopoly ISP facing homogeneous consumers with inelastic subscription. In addition, we find that that ISP's total profit increases with the profit from the content side. This result implies that ISPs will make an interconnection agreement and pricing decisions to maximize their profits from the content side, and thus simplifies the analysis of the two-sided markets considerably. Sections 6 analyzes ISPs' choice of qualtiy and access charges in each regime and compare them. In particular, we derive conditions under which "bill and keep" arises endogenously as an euqilibrium outcome. Section 7 conducts a welfare analysis and derives conditions under which the neutral regime can outperform the non-neutral one in terms of social welfare. This result shows the importance of CPs' business models in the evaluation of net neutrality regulation. Section 8 contains concluding remarks, along with suggestions for further possible extensions of our analysis.

2 A Model of Interconnected Networks with Competition

2.1 ISPs, CPs, and Consumers

We consider two interconnected ISPs denoted by $i = 1, 2$. ISPs serve as platforms in a two-sided market where CPs and end consumers constitute two distinct groups of customers. As pointed out by LMRT (2001, 2003), the traffic between CPs and the traffic between consumers such as E-mail exchanges take up trivial volumes relative to the volume of traffic from CPs to consumers. Thus, we focus on the primary traffic from CPs to consumers who browse web pages, download files, stream multi-media content, etc. As in the standard literature of interconnected networks, we assume a *balanced traffic pattern* that consumers' interest in a CP is independent of the CP's ISP choice and *reciprocal access pricing* that implies no asymmetry in the access charge for incoming and outgoing traffics.

There is a continuum of CPs whose mass is normalized to one. We consider a very simple case of CP heterogeneity. There are two types of CPs: $\theta \in \{\theta_H, \theta_L\}$, with $\theta_\Delta = \theta_H - \theta_L > 0$. The measure of θ_k type CP is denoted by ν_k, where $k = H, L$, and $\nu_H = \nu$ and $\nu_L = 1 - \nu$. There is also a continuum of consumers who demand one unit of each content whose value depends on content type θ and its quality q. In our context, quality means speed and reliability of content delivery. Let q_k denote the quality of delivery associated with content of type θ_k. The total surplus generated from interaction between a consumer and a CP of type θ is equal to $\theta u(q)$, where $u' > 0$ and $u'' < 0$. According to our utility formulation, the parameter θ reflects the sensitivity of content to delay, with higher valuation content being more time/congestion sensitive. Note that $\theta u(q)$ captures not only a consumer's gross surplus but also a CP's revenue from advertising. We assume that this surplus is divided between a CP and a consumer such that the former gets $\alpha\theta u(q)$ and the latter $(1-\alpha)\theta u(q)$ with $\alpha \in [0,1]$. The parameter α reflects the nature of CPs' business model. We have in mind two sources of revenue for CPs: micropayments and advertising revenue.[3] The parameter α is higher when CPs use both instruments than when CPs use only micropayments, which in turn is higher than when CPs make money only from advertising.

The two ISPs are horizontally differentiated in the consumer side. In order to model elastic participation of consumers, we adopt a "Hotelling model with hinterlands" as in Armstrong and Wright (2009) and Hagiu and Lee (2011). Specifically, the demand for network i is given by

$$n_i = \frac{1}{2} + \frac{U_i - U_j}{2t} + \lambda U_i, \tag{1}$$

where n_i is the measure of consumers subscribing to ISP i, t is the "transportation cost" parameter, and $\lambda \geq 0$ is a parameter representing the relative importance of market expansion possibilities. A consumer's net utility from subscribing to ISP i (gross of the transportation cost), U_i, is given as follows.

$$U_i(\alpha, \mathbf{q}, u_0, f_i) = u_0 + (1-\alpha) \sum_{\theta_k \in \Theta_i} \nu_k[\theta_k u(q_k)] - f_i \tag{2}$$

where u_0 is the intrinsic utility associated with the Internet connection and f_i is the subscription price charged by ISP i. The type space Θ_i is the set of CP types served by ISP i. Assume that u_0 is large enough relative to t such that the consumer market is always covered for $\lambda = 0$.

Concerning the market expansion possibilities, $\lambda = 0$ corresponds to the standard Hotelling model with inelastic subscription in which consumers of mass one are uniformly located on the Hotelling line and the two ISPs are located at the end points. $\lambda > 0$ corresponds to a situation in which each ISP faces a downward sloping demand of loyal consumers on each side of the unit interval that can be considered as "hinterlands" of each ISP. Consumers in these areas never consider buying

[3] Our formulation captures the idea that higher valuation content can generate higher advertising revenues.

from the alternative ISP. These consumers also have the same transportation cost parameter of t, but are uniformly distributed with a density of H. In such a scenario, the marginal consumer type in the loyal consumer group who is indifferent between not subscribing and subscribing to ISP i is distance x_i^* away (in the hinterlands) from the ISP, where x_i^* is defined by

$$U_i - tx_i^* = 0.$$

Thus, the number of consumers in the hinterland is given by $Hx_i^* = \lambda U_i$ with $\lambda = \frac{H}{t}$.[4]

2.2 Network Interconnection and Network Neutrality

ISPs provide network services that deliver content from CPs to consumers. The marginal cost of providing a unit traffic of quality q from CP to end users is assumed to be linear, i.e., $c(q) = cq$ for $q \geq 0$.[5] Note that $c = c_O + c_T$, where $c_O \geq 0$ and $c_T \geq 0$ stand for the cost of origination and that of termination per quality, respectively. We assume that ISPs cannot engage in first-degree price discrimination across content providers depending on content types.

We consider two different regimes under which ISPs can deliver content. For the simplicity of analysis, we consider cooperative choice of quality and access charge for both neutral and non neutral networks. Under a non-neutral regime, ISPs can offer multiple classes of services that differ in delivery quality. In other words, they can engage in second degree price discrimination by offering a menu of contracts that charges different prices depending on the quality of delivery. Let q_H be the quality for high type CPs and q_L the one for low type CPs; let A_H and A_L denote the reciprocal termination charges for each quality class. Let us define $a_k = \frac{A_k}{q_k}$. Then, for one unit of off-net traffic of quality $q = q_H$ from ISP j to ISP i (i.e., a consumer subscribed to ISP i asks for content from a CP subscribed to ISP j), the origination ISP j incurs a cost of $c_O q_H$ and pays an access charge of $a_H q_H$ to ISP i, and the termination ISP i incurs a cost of $c_T q_H$ and receives an access charge of $a_H q_H$ from ISP i. Let $\widehat{c}_k \equiv c + a_k - c_T$ $(= c_O + a_k)$ denote the perceived unit quality cost of the off-net content that terminates in the other network for $q = q_k$, where $k = H, L$.

In a non-neutral regime, we further distinguish two cases depending on whether or not termination-based price discrimination (TPD) is possible. With TPD, ISP i proposes a non-linear pricing

[4] The market expansion possibility parameter λ can also be represented by the same density of consumers in the hinterlands (i.e, $H = 1$), but with a different transportation parameter for consumers in the hinterlands, say t_H.

[5] The assumption of a linear marginal cost in quality can be made without any loss of generality because we can normalize quality to satisfy the assumption of linearity. Suppose that $c(q)$ is nonlinear. By redefining $\widetilde{q}$ as $c(q)/c$, we have a linear marginal cost function $\widetilde{c}(\widetilde{q}) = c\widetilde{q}$. Starting from a concave utility function and a convex cost function, after this linealization, the utility function with the normalized cost remains still concave.

schedule $\{p_i(q), \widehat{p}_i(q)\}$ for $q \in \{q_H, q_L\}$ such that upon paying $p_i(q)$ (respectively, $\widehat{p}_i(q)$) a CP can obtain delivery of its content with quality q from ISP i for a unit of on-net traffic (respectively, a unit of off-net traffic). No TPD is a particular case of TPD in which $p_i(q) = \widehat{p}_i(q)$.

In a neutral regime or in the presence of net neutrality regulation, ISP i is constrained to offer a single uniform delivery quality q cooperatively with the corresponding price p. There is no TPD because content cannot be treated differently depending on its destination. The ISPs jointly choose a single quality level and a single access charge a . Let $\widehat{c} \equiv c + a - c_T$ denote the off-net cost per unit quality in the case of neutral networks. Define $\widehat{\mathbf{c}}$ such that $\widehat{\mathbf{c}} \equiv (\widehat{c}_H, \widehat{c}_L)$ in the case of non-neutral networks and $\widehat{\mathbf{c}} \equiv \widehat{c}$ in the case of neutral networks.

Figure 1 illustrates the flows of traffic and payment over two interconnected networks.

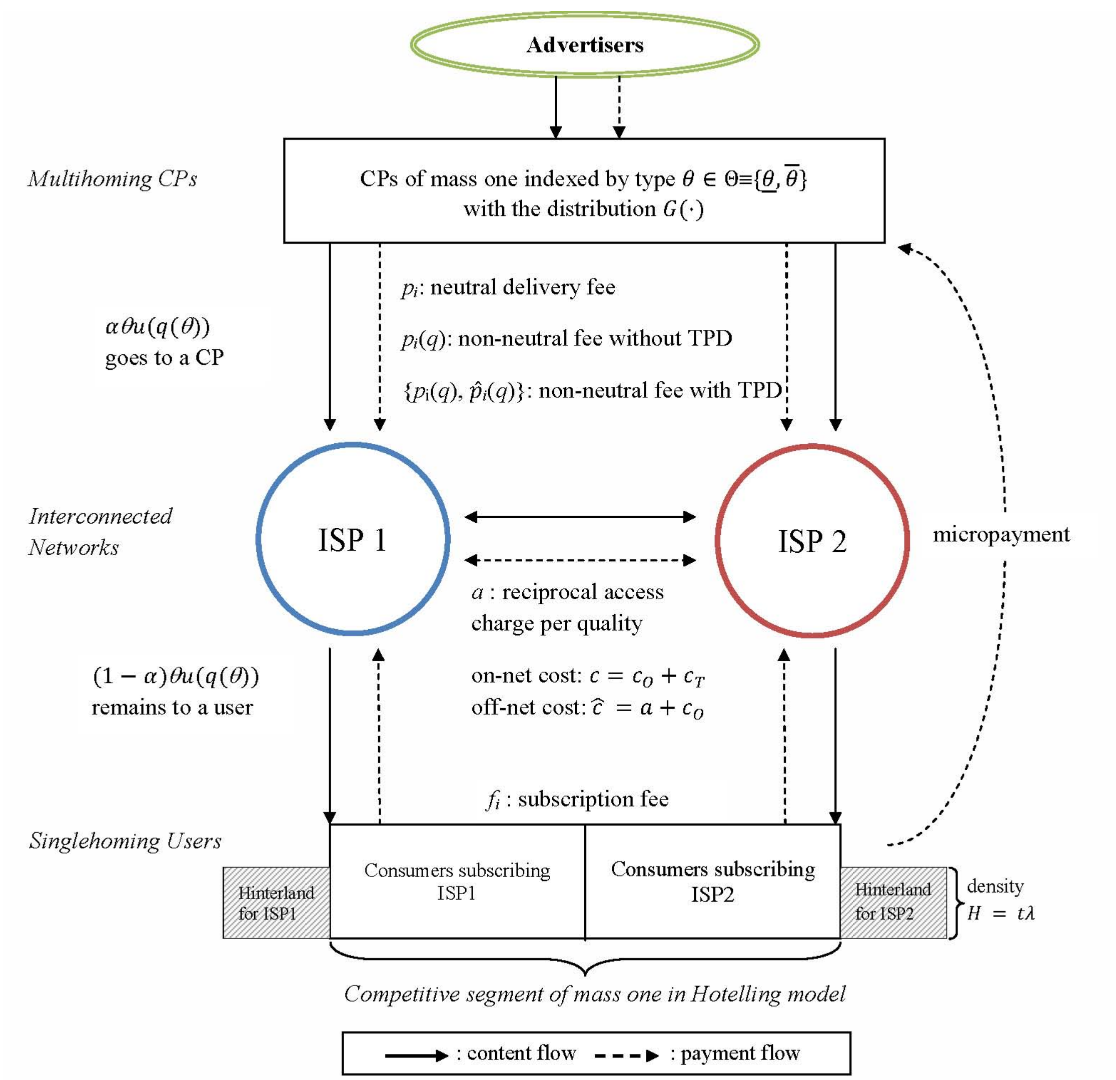

Figure 1: The flows of traffic and payment over interconnected networks

2.3 Timing of Decisions

1. The quality levels and the corresponding access charges are negotiated between the ISPs.

2. Each ISP i with $i = 1, 2$ simultaneously determines delivery fees for CPs for both on-net and off-net traffics, $\{p_i(q_H), p_i(q_L)\}$, in the case of no neutrality without TPD. (Corresponding price(s) will be determined for other cases.) CPs can multihome and decide which ISPs to use to deliver their content.

3. Each ISP i with $i = 1, 2$ simultaneously posts its consumer subscription fee f_i and consumers single-home and make their subscription decisions.

The reason why we consider this sequential timing is two-fold. First, the sequential timing with reverse order of 2 and 3 creates a serious commitment problem. Once consumers paid for the subscription price, then each ISP has no incentive to subsidize CPs. Second, we find that the current timing has advantage over the simultaneous pricing game when we consider ISPs' deviation incentives. Consider an equilibrium price under the simultaneous timing. If ISP i deviates in terms of its offer to CPs, it can also adjust its offer to consumers but ISP j cannot. In contrast, in the sequential timing that we consider, if ISP i deviates in stage 2 by changing the match between quality and CP type, ISP j can adjust its offer to consumers in stage 3. For this reason, the current timing minimizes the incentives to deviate.

3 Benchmarks

In this section, we consider two benchmarks: the first-best and a monopoly ISP facing homogeneous consumers. The second benchmark is used to introduce the key assumptions of our paper.

3.1 First-Best

Before analyzing market outcomes under various regimes, we first analyze the first-best outcome as a benchmark. For any given configuration of consumers subscribing to the networks, the socially optimal quality level for CPs of type θ_k, denoted by q_k^{FB}, is determined by the following condition.

$$\theta_k u'(q_k^{FB}) = c, \text{ where } k = H, L \tag{3}$$

The marginal benefit of an incremental improvement of delivery quality for the content of type θ must be equal to c, the marginal cost associated with such an adjustment.

Define u^{FB} and c^{FB} as the gross utility from all content providers and its associated content delivery cost for each consumer when the first best delivery qualities are chosen.

$$\begin{aligned} u^{FB} &= \sum_{k=H,L} \nu_k \theta_k u(q_k^{FB}) \\ c^{FB} &= \sum_{k=H,L} \nu_k c q_k^{FB} \end{aligned}$$

Then, social welfare as a function of the measure of subscribed consumers, denoted by $N(\geq 1)$, is given by

$$SW(N) = N \times \left(u_0 + u^{FB} - c^{FB}\right) - T(N),$$

where $T(N)$ represents the total transportation cost incurred by consumers. Since the number of consumers in the competitive market is normalized to one, the total number of consumers from the hinterlands is given by $(N-1)$. The total transportation cost $T(N)$ is minimized with a symmetric subscription pattern to the two ISPs. Let x be the distance from the marginal consumer to the closest ISP with $(N-1)/2$ subscribers. Then, $x = \frac{N-1}{2H}$. With our "hinterlands" specification, we thus have

$$T(N) = 2\int_0^{\frac{1}{2}} tx dx + 2H \int_0^{\frac{N-1}{2H}} tx dx = \frac{t}{4} + \frac{(N-1)^2}{4\lambda}$$

The first-best measure of subscribed consumers, denoted by N^{FB}, is given by the following first order condition:

$$u_0 + u^{FB} - c^{FB} = \left.\frac{dT(N)}{dN}\right|_{N=N^{FB}} = \frac{N^{FB}-1}{2\lambda} (= tx^{FB}) \quad (4)$$

The transportation cost of the marginal consumer (tx^{FB}) should be equal to $u_0 + u^{FB} - c^{FB}$ in the first-best outcome.

Proposition 1 *(First-best) The first best outcome requires no neutrality and the first-best quality schedule q_k^{FB}, $k = H, L$, is given by (3). Under perfect price discrimination, the first-best outcome can be implemented by a price schedule to CPs $p(q_k^{FB}) = \alpha\theta u(q_k^{FB})$ and consumer subscription price $f = c^{FB} - \alpha u^{FB}$. It requires subsidy on CPs of type θ_k if $cq_k^{FB} > \alpha\theta_k u(q_k^{FB})$ and a subsidy on consumer side if $c^{FB} < \alpha u^{FB}$. Then, each CP and each ISP realize zero profit.*

The point of the proposition is simple. With heterogeneous content that differs in sensitivity to delivery quality, the uniform treatment of content mandated by net neutrality in general would not yield a socially optimal outcome. The prices described in Proposition 1 are the unique ones that implement the first-best outcome under the budget constraint of the social planner: each ISP and each CP realizes zero profit and the marginal consumer is indifferent between subscribing and not subscribing.

3.2 Monopoly ISP with Homogeneous Consumers and Key Assumptions

As another benchmark, we consider a hypothetical setting in which a monopoly ISP provides content delivery service from two types of CPs to *homogeneous* consumers. This benchmark serves a crucial role in characterizing the equilibrium with competing ISPs because we establish an equivalence result between the monopolistic outcome and the competitive outcome. This benchmark is also useful in introducing our key assumptions. We normalize, without loss of generality, the total measure of subscribed consumers to one. We assume that the monopoly simultaneously announces the fee for consumers and the price-quality pairs for CPs.[6]

■ Non-neutral Networks

Let $\{(p_H, q_H), (p_L, q_L)\}$ be the menu of cotract offered to CPs. Then, each consumer's gross utility is given by $U(\alpha) = u_0 + (1-\alpha)\sum_{k=H,L} \nu_k[\theta_k u(q_k)]$, which can be fully extracted by a subscription fee f. The ISP's profit from the content side is $\pi^{CP}(\alpha) = \sum_{k=H,L} \nu_k\,[p_k - cq_k]$. The overall profit for the ISP can be written as $\Pi(\alpha) = U(\alpha) + \pi^{CP}(\alpha)$. Thus, the monopolistic ISP's mechanism design problem can be described as:

$$\max_{(p_j,q_k)} \Pi(\alpha) = \sum_{k=H,L} \nu_k[p_k + (1-\alpha)\theta_k u(q_k) - cq_k] + u_0$$

subject to

$$\begin{aligned}
IC_H &: \alpha\theta_H u(q_H) - p_H \geq \alpha\theta_H u(q_L) - p_L \\
IC_L &: \alpha\theta_L u(q_L) - p_L \geq \alpha\theta_L u(q_H) - p_H \\
IR_H &: \alpha\theta_H u(q_H) - p_H \geq 0 \\
IR_L &: \alpha\theta_L u(q_L) - p_L \geq 0.
\end{aligned}$$

[6]Equivalently, we can assume that the monopolist first chooses the price-quality pairs and then announce the fee for consumers.

This is a standard mechanism design problem. As usual, the high-type's incentive compatibility constraint IC_H and the low-type's individual rationality constraint IR_L are binding: we have

$$p_H = \alpha\theta_H u(q_H) - \alpha\theta_\Delta u(q_L); \qquad p_L = \alpha\theta_L u(q_L). \tag{5}$$

This leads to the following reduced problem

$$\max_{\{q_H, q_L\}} \quad \Pi(\alpha) = u_0 + \sum_{k=H,L} \nu_k[\theta_k u(q_k) - cq_k] - \nu \cdot \alpha\theta_\Delta u(q_L).$$

The objective in the reduced program shows that the ISPs extract full surplus except for the informational rent to high type CPs, which is given by $\nu \cdot \alpha\theta_\Delta u(q_L)$. From the first order conditions, we find that the optimal quality for the high type is determined by $\theta_H u'(q_H^*) = c$, which is equal to the first-best level: there is no quality distortion for the high type, regardless of α. On the contrary, the low type CPs' quality is characterized by

$$\left(\theta_L - \frac{\nu}{(1-\nu)} \cdot \alpha\theta_\Delta\right) u'(q_L^*(\alpha)) = c. \tag{6}$$

Our first assumption is that the monopoly ISP prefers serving both types when it can practice second-degree price discrimination when CPs extract all the surplus, that is, α=1.

Assumption 1 $q_L^*(\alpha = 1) > 0$

$q_L^*(\alpha = 1)$ requires $\theta_L > \frac{\nu}{1-\nu}\theta_\Delta$. Assumption 1 ensures that $q_L^*(\alpha) > 0$ for any $\alpha \in [0, 1]$ because total differentiation applied to (6) shows that the low-type quality is decreasing in α :

$$\frac{dq_L^*}{d\alpha} = \frac{\nu\theta_\Delta u'(q_L^*)}{((1-\nu)\theta_L - \nu\alpha\theta_\Delta)\, u''(q_L^*)} < 0 \tag{7}$$

This is because the ISP has the stronger incentive to distort the quality for the low type as the greater surplus is extracted from the CP side. So, the quality for the low-type CP does not reach the first-best quality unless $\alpha = 0$: $q_L^{SB} \leq q_L^{FB}$ follows from (6) and $u''(\cdot) < 0$ and the equality holds only if $\alpha = 0$.

■ Neutral Networks

Now consider a neutral network where the ISP is constrained to choose only a single price-quality pair (p, q). The ISPs can decide between excluding low type CPs and serving both types of CPs.

With the exclusion, it is straightforward that the ISP will choose $q = q_H^{FB}$ and $p = \alpha\theta_H u(q_H^{FB})$, which gives $\Pi^{EX} = \nu[\theta_H u(q_H^{FB}) - cq_H^{FB}] + u_0$. Note that this profit is independent of α. If the monopolistic ISP decides to serve both types, then $p = \alpha\theta_L u(q)$ and $f = u_0 + (1-\alpha)\sum_{k=H,L}\nu_k\theta_k u(q)$. Hence, the monopoly ISP chooses a single quality q to solve

$$\max_q \ \Pi(\alpha) = u_0 + (\theta_L + (1-\alpha)\nu\theta_\Delta)u(q) - cq.$$

From the first-order condition, we obtain:

$$(\theta_L + (1-\alpha)\nu\theta_\Delta)u'(\widetilde{q}(\alpha)) = c \tag{8}$$

Totally differentiating (8) shows that the quality decreases with α:

$$\frac{d\widetilde{q}(\alpha)}{d\alpha} = \frac{\nu\theta_\Delta u'(\widetilde{q})}{(\theta_L + (1-\alpha)\nu\theta_\Delta)u''(\widetilde{q})} < 0. \tag{9}$$

From (8), it is straightforward that $\widetilde{q}(\alpha) \geq q_L^{FB}$ where the equality holds when $\alpha = 0$.

From the envelope theorem, the maximized objective without exclusion $\widetilde{\Pi}(\alpha) = u_0 + (\theta_L + (1-\alpha)\nu\theta_\Delta)u(\widetilde{q}(\alpha)) - c\widetilde{q}(\alpha)$ strictly decreases with α. On the contrary, the maximized objective under exclusion $\widetilde{\Pi}^{EX} = \nu\cdot(\theta_H u(q_H^{FB}) - cq_H^{FB})$ is constant. We assume the following.

Assumption 2 $\widetilde{\Pi}(\alpha = 0) > \widetilde{\Pi}^{EX} > \widetilde{\Pi}(\alpha = 1)$

The monotonicity $\widetilde{\Pi}(\alpha)$ then implies that there exists a unique threshold level of α denoted by $\alpha^N \in (0,1)$ where it is implicitly defined by $\widetilde{\Pi}(\alpha^N) = \widetilde{\Pi}^{EX}$, that is,

$$(\theta_L + (1-\alpha^N)\nu\theta_\Delta)u(\widetilde{q}(\alpha^N)) - c\widetilde{q}(\alpha^N) = \nu\cdot(\theta_H u(q_H^{FB}) - cq_H^{FB}). \tag{10}$$

No exclusion strategy yields higher profit to the ISP when α is less than α_N, whereas exclusion becomes more profitable than serving both types when α exceeds this thresold.

Therefore, the quality chosen by the ISP is given by $q^*(\alpha) = q_H^{FB}$ for $\alpha > \alpha^N$ and $q^*(\alpha) = \widetilde{q}(\alpha)$ otherwise. Accordingly, the retail price is given by $p^*(\alpha) = \alpha\theta_H u(q_H^{FB})$ for $\alpha > \alpha^N$ and $p^*(\alpha) = \alpha\theta_L u(\widetilde{q}(\alpha))$ otherwise.

■ Assumptions and Social Welfare

Two remarks on our assumptions are in order. First, Assumptions 1 and 2 correspond to a situation analyzed in Hermalin and Katz (2007). Specifically, when a CP extracts the entire surplus, the monopoly ISP prefers excluding the low-type CP without a second degree price discrimination, though it serves both types with it.

Second, from social welfare point of view, the non-neutral network dominates the neutral network for the extreme cases of $\alpha = 1$ and $\alpha = 0$. Essentially, these two cases can considered as a representation of one-sided markets. Consider first the case in which CPs capture the whole surplus from interactions with consumers, i.e., $\alpha = 1$. Then, each consumer obtains the basic utility u_0 only. So, the monopoly ISP will set $f = u_0$ both under non-neutral and neutral networks. Consequently, we can focus on the monopoly ISP's problem of maximizing profit from CPs, which is a standard problem of one-sided market. In this case, high type CPs consume q_H^{FB} in both regimes, but low types are served only under non-neutral network. This is a standard argument for second-degree price discrimination.

For the other extreme case of $\alpha = 0$, consumers capture all surplus from interactions with CPs. Since consumers are homogeneous, the monopoly ISP can extract full surplus from consumers. The case of $\alpha = 0$ is the same as a standard monopoly in one-sided market with cost function cq. The monopoly will provide services for free to CPs, which means that the ISP bears the entire cost of cq. When $\alpha = 0$, the monopoly ISP provides the first-best quality for each type of CPs and charges the consumer subscription fee $f(\alpha = 0) = u_0 + u^{FB}$. The ISP earns the profit of $\Pi(\alpha = 0) = u_0 + u^{FB} - c^{FB}$. Under a neutral network, again the ISP has two choices. First, it can allow both types to send their content for free. Then, the ISP chooses the quality q equal to q^* defined by $[\nu\theta_H + (1-\nu)\theta_L]\, u'(q^*) = c$ and earns the profit of $u_0 + [\nu\theta_H + (1-\nu)\theta_L]\, u(q^*) - cq^*$. Second, the ISP can serve only the high-type CPs for free (with the low-type excluded),[7] and earn the profit of $u_0 + \nu[\theta_H u(q_H^{FB}) - cq_H^{FB}]$. The monopoly ISP under neutral network prefers no exclusion to exclusion if $[\nu\theta_H + (1-\nu)\theta_L]\, u(q^*) - cq^* > \nu[\theta_H u(q_H^{FB}) - cq_H^{FB}]$ holds. By definition of the first-best, the ISP and a social planner prefer a non-neutral network over a neutral network for $\alpha = 0$.

Summarizing thus far analysis, we have:

Proposition 2 *(Monopoly ISP) Consider a monopoly ISP facing homogenous consumers with inelastic subscription.*

(a) In the case of $\alpha = 1$, under Assumptions 1-2, the ISP serves both types of CPs in a non-

[7]When $\alpha = 0$, every CP makes zero profit and we can assume that in the case of indifference, a CP follows the ISP' desire. For any $\alpha > 0$ (hence α can be as close as possible to zero) and $q > 0$, the ISP can exclude low types by charging $p = \alpha\theta_H u(q)$.

neutral network whereas it serves only high types in a neutral network. Therefore, social welfare is higher under a non-neutral network than under a neutral network.

(b) In the case of $\alpha = 0$, the outcome chosen by the ISP coincides with the first-best under a non-neutral network. Therefore, social welfare is higher under a non-neutral network than under a neutral network.

Under Assumptions 1-2, we consider a scenario in which the neutral network is always dominated in one-sided market settings, stacking the deck against the neutral network. Later, this result will be contrasted to the case where a neutral network can provide a higher social welfare relative to a non-neutral network.

4 Networks Competition in CP Market and Off-net Cost Pricing

In this section, we analyze the ISPs' competition in the content market and establish that any equilibrium prices for CPs should satisfy the *off-net cost pricing principle*, which we use for subsequent analysis. In the beginning of stage 2, quality levels and access charges are given from stage 1. Let N be the total number of consumers that will be determined in stage 3. We denote ISP i's consumer market share by $s_i \in [0, 1]$, that is, $s_i = \frac{n_i}{N}$.

First, under a neutral network (hence without termination-based price discrimination), consider an equilibrium price for CPs $p(q)$ given that the ISPs previously agreed on (q, a). Suppose first that at $p(q)$ both types buy connections from ISP i. Then, ISP i's profit from content side in equilibrium is $p(q) - s_i cq - (1 - s_i)\widehat{c}q$, where $\widehat{c} = (c_O + a)$ is the off-net marginal delivery cost per quality. If it loses the CPs by charging a higher price, its only source of profit from content side is via deliving off-net traffics which results in $s_i(a - c_T)q$. Therefore, the following inequality must hold in equilibrium:

$$p(q) - s_i cq - (1 - s_i)(c_O + a)\,q \geq s_i(a - c_T)q,$$

which is equivalent to

$$p(q) \geq (c_O + a)\,q = \widehat{c}q.$$

Symmetrically, the condition for ISP j to weakly prefer losing CPs to winning CPs gives the condition

$$p(q) \leq (c_O + a)\,q = \widehat{c}q.$$

Therefore, any equilibrium price should satisfy $p(q) = \widehat{c}q$.

Suppose now that at $p(q)$ only high types buy connections and ISP i wins them. Then, ISP i's equilibrium profit from content side is $\nu\left[p(q) - s_i cq - (1 - s_i)(c_O + a)q\right]$ and its profit from content side from losing CPs is $\nu s_i(a - c_T)q$. The previous logic still applies here again. Therefore, any equilibrium price should satisfy $p(q) = \widehat{c}q$ regardless of whether exclusion of low types occurs or not. The same result holds when we consider non-neutral networks without termination-based price discrimination.

Consider now a non-neutral network with termination-based price discrimination. More specifically, let us consider competition between ISP i's on-net offer and ISP j's off-net offer in the provision of access to ISP i's consumers. Let (q_H, a_H) and (q_L, a_L) be the quality-access charge pairs on which the ISPs previously agreed upon. Suppose that there is a CP that purchases delivery service of quality q_k at the price of p_k with $k = H, L$. If ISP i wins the business of the CP and sells the on-net traffic to this CP, it must hold that ISP i prefers providing the on-net traffic to delivering off-net traffics from the CP to its end users. This implies that

$$p_k - cq_k \geq (a_k - c_T)\, q_k$$

The inequality is equivalent to

$$p_k \geq (c_O + a_k)\, q_k = \widehat{c}_k q.$$

The ISP j prefers losing CPs to winning them if and only if

$$0 \geq p_k - (c_O + a_k)\, q_k$$

since the ISP j's cost supplying q_k is $(c_O + a_k)\, q_k$. Therefore, the off-net cost pricing must hold. The same logic applies to the case in which the other ISP wins CPs in equilibrium.

Summarizing the results, we have:

Lemma 1 *(Off-net cost pricing) Any equilibrium prices that generates positive sales to CPs must satisfy off-net cost pricing. This holds regardless of whether or not networks are neutral and whether or not there is TPD.*

The lemma shows that off-net cost pricing is a necessary condition that any equilibrium price for CPs generating positive sales must satisfy. However, off-net cost pricing is not a sufficient condition because an ISP may have an incentive to deviate. To illustrate this point, consider a neutral network. Suppose that $a = (\alpha\theta_L u(q) + \varepsilon)/q - c_O$ where $\varepsilon > 0$ is infinitesimal. Then, off-net cost pricing leads to $p(q) = \alpha\theta_L u(q) + \varepsilon$. Hence, only high types purchase the quality at off-net cost pricing. Consider now the deviation of ISP i to $p'(q) = \alpha\theta_L u(q)$ such that both types purchase the quality. This deviation is profitable iff

$$\nu s_i(\alpha\theta_L u(q) + \varepsilon - cq) < \alpha\theta_L u(q) - s_i cq - (1 - s_i)\left(\alpha\theta_L u(q) + \varepsilon\right),$$

which is equivalent to

$$\left[\nu s_i + (1 - s_i)\right]\varepsilon < (1 - \nu)s_i(\alpha\theta_L u(q) - cq),$$

which holds for $\varepsilon > 0$ small enough as long as $\alpha\theta_L u(q) > cq$.

From the discussions of this section, we have

Lemma 2 *(Profit from CPs) Consider any off-net cost pricing equilibrium. Assume that some type $\theta \in \{\theta_L, \theta_H\}$ buys some quality q at off-net cost $p(q) = (c + a - c_T)q$. Then, ISP i obtains a profit per CP of type θ equal to $n_i(a - c_T)q$ while ISP j realizes zero profit. This result holds regardless of whether networks are neutral or not and of whether there is termination based price discrimination.*

Note first that the result of this lemma does not depends on which ISP wins consumers since each ISP is indifferent between winning and losing CPs at off-net cost pricing. Also, note that each ISP realizes profit only from on-net market in the case of termination-based price discrimination. Lemma 2 allows us to write each ISP's profit from content side as $n_i\widehat{\pi}^{CP}$, where $\widehat{\pi}^{CP} = \pi^{CP}(p = \widehat{c}q) = \sum_{k=H,L} \nu_k\,(\widehat{c} - c)q_k$ denotes the profit *per consumer* that each ISP makes from the content side and does not depend on (n_1, n_2).

5 Networks Competition in Consumer Subscription Market

In the previous section, we showed that off-net cost pricing must be satisfied in any equilibrium and that under off-net cost pricing, each ISP's profit from content side is given by $n_i\widehat{\pi}^{CP}$. Given these, results, let us study the competition between two ISPs in the consumer subscription market. Therefore, ISP i's total profit is given by

$$\Pi_i = n_i \cdot (f_i + \widehat{\pi}^{CP}),\ \text{where } n_i = \frac{1}{2} + \frac{f_j - f_i}{2t} + \lambda U_i$$

In any interior equilibrium, each ISP i chooses f_i to maximize its total profit from both CPs and consumers, given f_j. Using the first order condition, symmetry, and the relationship $\lambda = \frac{H}{t}$, we derive the symmetric equilibrium subscription price:

$$f^*(\widehat{\pi}^{CP}; H) = \frac{t + 2HU(\alpha)}{1 + 4H} - \frac{(1 + 2H)\widehat{\pi}^{CP}}{1 + 4H}. \tag{11}$$

where $U(\alpha) = u_0 + (1 - \alpha) \sum_{k=H,L} \nu_k \theta_k u(q_k)$ is the consumer gross utility. So, the ISP's equilibrium profit is given by

$$\pi^*(\widehat{\mathbf{c}}; \alpha, t, \lambda) = \left(\frac{1}{2} + \lambda \left(U(\alpha) - f^* \right) \right) \cdot (f^* + \widehat{\pi}^{CP}).$$

where f^* is from (11). With some algebra, we find that

Proposition 3 *Consider any symmetric equilibrium.*

(i) Each ISP earns the profit

$$\pi^*(\widehat{\mathbf{c}}; \alpha, t, \lambda) = \left[\frac{1}{2} + \lambda \frac{(1 + 2H) \cdot \widehat{\Pi}(\alpha) - t}{1 + 4H} \right] \left[\frac{t + 2H \cdot \widehat{\Pi}(\alpha)}{1 + 4H} \right] \tag{12}$$

where $\widehat{\Pi}(\alpha) = U(\alpha) + \widehat{\pi}^{CP}$ *and* $\lambda = \frac{H}{t}$.

(ii) (a) For $H = t\lambda = 0$, *each ISP's profit is always equal to the Hotelling profit* $t/2$.

(ii) (b) For $H = t\lambda > 0$, *maximizing joint profit of the ISPs requires maximizing* $\widehat{\Pi}(\alpha) = U(\alpha) + \widehat{\pi}^{CP}$. *In other words, the ISPs maximizing joint profit behave as a monopoly ISP facing homogeneous consumers with inelastic subscription but constrained to engage in hypothetical off-net cost pricing* $p(q) = \widehat{c}q$.

The result of Proposition 3(ii)(b) is extremely interesting. Note first that for $H = t\lambda > 0$, the equilibrium profit (12) is an increasing function of $\widehat{\Pi}(\alpha)$. This implies that what the competing ISPs maximize when jointly choosing quality levels and access charges is equivalent to what a monopoly ISP maximizes when it faces homogenous consumers with inelastic subscription, which we analyzed as a benchmark in section 3.2. Recall that $\widehat{\Pi}(\alpha) \leq u_0 + u^{FB} - c^{FB}$, where the equality holds only when the ISPs capture the entire CPs' surplus with the first-degree price discrimination. This implies that there are potentially two sources of distortions in the objective of the ISPs with respect to social welfare: the ISPs neglect the rent of the CPs and endogenous subscription of consumers. Later, we will study how these two distortions play differently in non-neutral and neutral networks.

What drives our result is that for given $\left(U(\alpha), \widehat{\pi}^{CP}\right)$, the competition on consumer side leads to an equilibrium subscription fee of the form

$$f^* = \beta U(\alpha) - (1-\beta)\widehat{\pi}^{CP} + \frac{t}{1+4H}, \tag{13}$$

where $\beta = \frac{2H}{1+4H} \in (0, 1/2)$ for $H = t\lambda > 0$ and $\beta = 0$ for $H = t\lambda = 0$. Hence, when consumer subscription is inelastic ($\lambda = 0$), the equilibrium consumer subscription fee is given by $f^* = t - \widehat{\pi}^{CP}$ and hence each ISP obtains the standard Hotelling profit $t/2$. In this case, any profit from the content side is completely dissipated away in the consumer side. Imposing net neutrality or not has no impact on each ISP's profit. This is reminiscent of the *profit neutrality* result in competition of telecommunications networks (Laffont, Rey, Tirole, 1998a) where the access charge has no impact on the networks' profits in a Hotelling model (hence with inelastic consumer subscription). Our result is stronger in the sense that the profit depends neither on the level of access charge nor on the number of product lines ISPs are allowed to offer.

When consumer subscription is elastic with $H = t\lambda > 0$, the profit neutrality result no longer holds; there is partial pass-through of $\widehat{\pi}^{CP}$ into f. This is because as $\widehat{\pi}^{CP}$ increases, the consumer subscription price decreases but *less* than the change in $\widehat{\pi}^{CP}$ for any $H > 0$, obviously seen from

$$-1 < \frac{\partial f^*}{\partial \widehat{\pi}^{CP}} = -(1-\beta) < -\frac{1}{2}. \tag{14}$$

Inequality (14) in turn implies that each ISP's total profit per consumer increases with π^{CP} :

$$0 < \frac{\partial (f^* + \widehat{\pi}^{CP})}{\partial \widehat{\pi}^{CP}} = \beta < \frac{1}{2}.$$

More generally, an ISP's profit is equal to the number of consumers multiplied by the profit per consumer. From (13), profit per consumer $f^* + \widehat{\pi}^{CP}$ is given by $\beta\widehat{\Pi}(\alpha)$ and hence linearly increases with $\widehat{\Pi}(\alpha)$. Note that the number of consumers is given by $\frac{1}{2} + \lambda(1-\beta)\widehat{\Pi}(\alpha)$ from $\frac{1}{2} + \lambda\left(U(\alpha) - f^*\right)$, which also linearly increases with $\widehat{\Pi}(\alpha)$. Therefore, the ISPs will choose quality levels and access charges to maximize $\widehat{\Pi}(\alpha)$ and replicate the monopolistic solution derived in section 3.2.

6 ISPs' Choice of Quality and Access Charges

In this section, we analyze the ISPs' choice of quality levels and access charges. We showed that the competing ISPs maximize the same objective as a monopoly ISP facing homogenous consumers. We also showed that off-net cost pricing must hold in any equilibrium. However, note that not all off-net costs can be supported as equilibrium prices for CPs since an ISP might have an incentive to deviate from off-net cost pricing in stage 2.

We proceed in two steps. First, we consider a benchmark in which no ISP is allowed to deviate from the off-net cost pricing in stage 2. Therefore the ISPs behave the same way as the monopoly ISP facing homogenous consumers behaves in section 3.2. It is because there is one-to-one correspondence between the retail price of content delivery and the choice of access charge from the off-net cost pricing, $p(q_k) = \widehat{c}q = (c + a_k - c_T)q_k$. In other words, any second-best quality-price combinations that would be chosen by the monopolistic ISP can be replicated by agreeing to the same quality levels and appopriate choice of access fees. Essentially, both the monopolistic ISP and competing ISPs have the same objective function and the same instruments.

Second, we consider the original case in which any ISP is allowed to deviate from off-net cost pricing in stage 2. Because of this deviation possibility, some off-net costs cannot be sustained as equilibrium prices. Therefore, the set of outcomes that the ISPs can achieve in this original case are contained in the set of outcomes that the ISPs can achieve without any deviation. This implies that the maximum joint profit that the ISPs can achieve without any deviation is an upper bound of the joint profit that they can achieve in the original case. We will check the sufficient conditions that allow the ISPs to achieve the upper bound.

6.1 Non-Neutral Networks

We now consider ISPs' choice of quality levels and access charges in the case of non-neutral networks for $H = t\lambda > 0$. According to Proposition 3(ii)(b), the ISPs maximize $\widehat{\Pi}(\alpha)$. Furthermore, according to Proposition 2, maximizing $\widehat{\Pi}(\alpha)$ under non-neutral networks requires to serve both type of CPs (i.e. no exclusion). Given the quality pairs (q_H, q_L) which the ISPs agreed to offer to each type of CPs, in the benchmark no ISP is allowed to deviate from the off-net cost pricing. The ISPs can indirectly choose the equilibrium price of each quality by properly choosing the access charge regardless of whether there is termination-based price discrimination.

From (5) and off-net cost pricing, it is immediate that the access charges will be chosen as follows to replicate the monopolistic solution:

$$a_H^* = \alpha\left(\theta_H u(q_H^*) - \theta_\Delta u(q_L^*)\right)/q_H^* - c_O \quad \text{and} \quad a_L^* = \alpha\theta_L u(q_L^*)/q_L^* - c_O. \tag{15}$$

Proposition 4 *Consider non-neutral networks and suppose that no ISP is allowed to deviate from the off-net cost pricing. Under Assumption 1:*

(i) The ISPs offer quality levels (q_H^*, q_L^*) *such that* $q_H^* = q_H^{FB}$ *for any* $\alpha \in [0,1]$ *and* $q_L^*(\alpha)$ *is determined by (6):* $q_L^*(0) = q_L^{FB}$ *and* $q_L^*(\alpha)$ *strictly decreases with* α*.*

(ii) The ISPs choose access charges (a_H^*, a_L^*) *given by (15). This leads to the following retail prices for CPs:* $p_H^* = \alpha\theta_H u(q_H^*) - \alpha\theta_\Delta u(q_L^*(\alpha))$ *and* $p_L^* = \alpha\theta_L u(q_L^*(\alpha))$*.*

Once we found the upper bound of the joint profits that the ISPs can realize, we now study under which conditions they can achieve the upper bound. In other words, we study under which condition no ISP has an incentive to deviate from off-net cost pricing. We normalize the total number of consumers subscribed at one without loss of generality.

Let us consider the case without TPD. Let $(\overline{q}, \underline{q})$ represent the quality allocated to high and low type. Note that ISP i is indifferent between winning CPs of a given type and losing them. Therefore, we need to consider only two deviation possibilities: ISP i can deviate to induce both types to buy $q_L^*(\alpha)$ or to buy q_H^*.

Consider first the deviation of ISP i to induce both types to consume low quality: i.e. $(\overline{q}, \underline{q}) = (q_L^*(\alpha), q_L^*(\alpha))$. Since the high type CPs are indifferent between the two qualities, ISP i can achieve this deviation at an epsilon discount of price. So the price it charges after the deviation is $p_i(q_L^*(\alpha)) = \alpha\theta_L u(q_L^*(\alpha))$. Since a_L^* the two ISPs agreed on in stage 1 is $a_L^* = \alpha\theta_L u(q_L^*(\alpha))/q_L^*(\alpha) - c_O$, we have off-net cost pricing; $p_i(q_L^*(\alpha)) = (a_L^* + c - c_T)q_L^*(\alpha)$. This implies that the stage 3 competition after the deviation leads to a symmetric equilibrium in which $(\overline{q}, \underline{q}) = (q_L^*(\alpha), q_L^*(\alpha))$ and $\pi_i^{CP} = \pi_j^{CP} = \pi^{CP} = (a_L^* - c_T)q_L^*(\alpha)$. This cannot give a higher profit than the upper bound that each ISP can obtain without deviation; otherwise, we have a contradiction to the upper bound.

Consider now the deviation of ISP i to induce both types to consume high quality: i.e. $(\overline{q}, \underline{q}) = (q_H^*, q_H^*)$. This requires ISP i to charge $p_i(q_H^*) = \alpha\theta_L u(q_H^*)$. Let (N, s_i, s_j) represent the total number of consumers subscribed and each ISP's consumer market share in stage 3. Then, ISP i's profit from content side is

$$\begin{aligned} & N\left[\alpha\theta_L u(q_H^*) - s_i c q_H^* - (1-s_i)(c + a_H^* - c_T)q_H^*\right] \\ = \; & N\left[\alpha\theta_L u(q_H^*) - (c + a_H^* - c_T)q_H^* + s_i(a_H^* - c_T)q_H^*\right] \\ = \; & N\left[-\alpha\theta_\Delta(u(q_H^*) - u(q_L^*(\alpha))) + s_i(a_H^* - c_T)q_H^*\right], \end{aligned}$$

where we use

$$a_H^* = \alpha\left[\theta_H u(q_H^*) - \theta_\Delta u(q_L^*(\alpha))\right]/q_H^* - c_O.$$

and ISP j's profit from content side is

$$N s_j (a_H^* - c_T) q_H^*.$$

The proof is based on the following conjecture.

- Digression to the conjecture

Fix $(\overline{q}, \underline{q}) = (q, q)$. Suppose that $(\overline{q}, \underline{q}) = (q, q)$ is implemented with off-net cost pricing such that it generates $\pi_i^{CP} = \pi_j^{CP} = \pi^{CP}$:

$$\begin{aligned} (a - c_T)q &= \pi^{CP}; \\ p(q) &= (c + a - c_T)q = \alpha\theta_L u(q). \end{aligned}$$

Consider now an asymmetric situation with a new access charge $a' = a + \Delta a$ with $\Delta a > 0$ in which ISP i is assumed to win all cps with the same retail price

$$p(q) = \alpha\theta_L u(q).$$

Hence, isp i's profit from content side is

$$N\left[\alpha\theta_L u(q) - s_i c q - (1-s_i)(c + a' - c_T)q\right];$$

isp j's one is

$$N s_j (a' - c_T) q = N s_j \left[(a - c_T)q + \Delta a q\right]. \tag{16}$$

Note

$$(c + a' - \Delta a - c_T)q = \alpha\theta_L u(q).$$

Hence, ISP i's profit from content side is

$$N\left[-\Delta a q + s_i(a' - c_T)q\right] = N\left[-(1 - s_i)\Delta a q + s_i(a - c_T)q\right]. \tag{17}$$

Note that a' will affect (N, s_i, s_j) determined in stage 3.

I have the following conjecture

Conjecture 1 *Consider the game defined before. Isp i's total profit is higher when* $\Delta a = 0$ *than when* $\Delta a > 0$.

INTUITION FOR THE CONJECTURE Basically, compared to when $\Delta a = 0$, what happens when $\Delta a > 0$ is such that on the one hand, from (17), it is as if isp i loses $\Delta a q$ per consumer susbcribed to the rival isp; on the other hand, from (16), it is as if isp j gains $\Delta a q$ extra per consumer susbcribed to isp j.

- Back to the proof

Consider $(\overline{q}, \underline{q}) = (q_H^*, q_H^*)$, $\Delta a q_H^* = \alpha\theta_\Delta(u(q_H^*) - u(q_L^*(\alpha)))$ and $a' = a_H^*$. Hence, we have

$$\begin{aligned}
(a - c_T)\, q_H^* &= (a_H^* - c_T - \Delta a)\, q_H^* \\
&= \alpha\left[\theta_H u(q_H^*) - \theta_\Delta u(q_L^*(\alpha))\right] - c_O q_H^* - \alpha\theta_\Delta\left[u(q_H^*) - u(q_L^*(\alpha))\right] \\
&= \alpha\theta_L u(q_H^*) - c_O q_H^*.
\end{aligned}$$

With the access charge a given by $(a - c_T)\, q_H^* = \alpha\theta_L u(q_H^*) - c_O q_H^*$, the off-net cost pricing leads to

$$(c + a - c_T)\, q_H^* = \alpha\theta_L u(q_H^*).$$

So from conjecture, the total profit of isp i upon deviation is smaller than the profit it obtains in a symmetric equilibrium with $(\overline{q}, \underline{q}) = (q_H^*, q_H^*)$ and a satisfying $(c + a - c_T)\, q_H^* = \alpha\theta_L u(q_H^*)$. Furthermore, the profit in this symmetric equilibrium is what the isps could achieve through off-net cost pricing and should give each isp a profit smaller than the upper bound. This ends the proof. Therefore, there is no profitable deviation from the upper bound of the joint profits characterized in 4.

6.2 Neutral networks

As in the non-neutral network, the monopolisitc ISP solution can be replicated by an appropriate chocie of the access charge if they are not allowed to deviate from the off-net cost pricing. More specifically, the ISPs will serve only high type CPs for $\alpha > \alpha^N$, and they will cooperatively choose the delivery quality level of $q^*(\alpha) = q_H^{FB}$ and the access charge of $a^*(\alpha) = \alpha\theta_H u(q_H^{FB})/q_H^{FB} - c_O$ to replicate the monopolistic solution. For $\alpha < \alpha^N$, the ISPs choose to serve both types of CPs with $q^*(\alpha) = \widetilde{q}(\alpha)$ and the corresponding access charge of $a^*(\alpha) = \alpha\theta_L u(\widetilde{q}(\alpha))/\widetilde{q}(\alpha) - c_O$.

Proposition 5 *Consider neutral networks and suppose that no ISP is allowed to deviate from the off-net cost pricing. Under Assumptions 1 and 2, there exists a unique threshold level of* α, *denoted by* $\alpha^N \in (0,1)$.

(i) The ISPs offer $q^*(\alpha) = q_H^{FB}$ *for* $\alpha > \alpha^N$ *and* $q^*(\alpha) = \widetilde{q}(\alpha)$ *otherwise.* $\widetilde{q}(\alpha)$ *is determined by (8) and it is higher than* q_L^{FB} *and strictly decreases with* α. *The cut-off value* α^N *is determined by defined by (10). The ISPs serve only high type CPs for* $\alpha > \alpha^N$ *and serve both types otherwise.*

(ii) The ISPs choose an access charge $a^*(\alpha) = \alpha\theta_H u(q_H^{FB})/q_H^{FB} - c_O$ *for* $\alpha > \alpha^N$ *and* $a^*(\alpha) = \alpha\underline{\theta}u(\widetilde{q}(\alpha))/\widetilde{q}(\alpha) - c_O$ *otherwise. This generates a retail price for CPs such that* $p^*(\alpha) = \alpha\theta_H u(q_H^{FB})$ *for* $\alpha > \alpha^N$ *and* $p^*(\alpha) = \alpha\theta_L u(\widetilde{q}(\alpha))$ *otherwise.*

The above result is intuitive. As the total size at stakes get smaller in the consumer side, the ISPs pay less attention to provide all types of CPs. Instead, they suffer more from the information rent by serving both types. Thus, exclusion strategy gets attractive as α increases. This finding has important policy implications. Recall that the parameter α may capture how much CPs can extract consumer surplus through micropayments. From this perspective, the concern about potential exclusion in a neutral network increases for the CPs whose business models are intensively based on some direct charge on consumers, not on advertising revenues with free access.

Corollary 1 $\frac{\partial \alpha^N}{\partial c} < 0$ *and* $\frac{\partial \alpha^N}{\partial \nu} = \frac{(1-\alpha^N)\Delta u(q^*) - (\theta_H u(q_H^{FB}) - cq_H^{FB})}{\nu \Delta u(q^*)} = \frac{cq^* - \theta_L u(q^*)}{\nu^2 \Delta u(q^*)} < 0$

As Corollary 1 shows, exclusion strategy is more likely to occur when the marginal cost increases and the proportion of high-type CP increase, other things being equal. Thus, the non-neutral network is likely to increase the allocative efficiency when heavy-bandwidth content are more prevalent in the Internet, which justifies the concern that equal treatment of all types of content may not be efficient way of using scarce network capacity.

We below show that the upper bound of the joint profits for neutral networks can be achieved when any ISP is allowed to deviate from the off-net cost pricing.

First, suppose that no type is excluded in the upper bound $(q_H, q_L) = (\widetilde{q}, \widetilde{q})$. Then it is clear that there is no profitable deviation. Increasing price for cps by isp i attracts no cps and hence does not affect π_i^{CP}, π_j^{CP}. Reducing price only reduces π_i^{CP}.

Second, consider the case in which low type is excluded $(q_H, q_L) = (q_H^{FB}, 0)$. More precisely, suppose that the two ISPs agreed on providing quality q_H^{FB} at access charge $a^* = \alpha\theta_H u(q_H^{FB})/q_H^{FB} - c_O$. Then, off-net cost pricing leads to $p^*(q_H^{FB}) = \alpha\theta_H u(q_H^{FB})$ and each ISP i realizes a profit of $\nu s_i(a^* - c_T)q_H^{FB}$.

The previous argument can be applied to show that there is no profitable deviation conditional on that only high type is served. Hence, it is enough to consider ISP i's deviation to serve both types such that $(q_H, q_L) = (q_H^{FB}, q_H^{FB})$; then it will choose $p_i(q_H^{FB}) = \alpha\theta_L u(q_H^{FB})$ and obtain a profit of

$$\begin{aligned} & N\left[\alpha\theta_L u(q_H^{FB}) - s_i c q_H^{FB} - (1-s_i)\left(c + a^* - c_T\right) q_H^{FB}\right] \\ = \; & N\left[-\alpha\Delta\theta u(q_H^{FB}) + s_i\left(a^* - c_T\right) q_H^{FB}\right] \end{aligned}$$

Isp j's profit is

$$N\left[s_j\left(a^* - c_T\right) q_H^{FB}\right].$$

Hence, we can apply the previous conjecture. Consider $(\overline{q}, \underline{q}) = (q_H^{FB}, q_H^{FB})$, $\Delta a q_H^{FB} = \alpha\Delta\theta u(q_H^{FB})$ and $a' = a^*$. Hence, we have

$$\begin{aligned} \left(a - c_T\right) q_H^* &= \left(a^* - c_T - \Delta a\right) q_H^{FB} \\ &= \alpha\theta_H u(q_H^{FB}) - c_O q_H^{FB} - \alpha\Delta\theta u(q_H^{FB}) \\ &= \alpha\theta_L u(q_H^{FB}) - c_O q_H^{FB}. \end{aligned}$$

With the access charge a given above, the off-net cost pricing leads to

$$\left(c + a - c_T\right) q_H^{FB} = \alpha\theta_L u(q_H^{FB}).$$

So from the conjecture, the total profit of isp i upon deviation is smaller than the profit it obtains in a symmetric equilibrium with $(\overline{q}, \underline{q}) = (q_H^{FB}, q_H^{FB})$ and a satisfying $(c + a - c_T)\, q_H^{FB} = \alpha \theta_L u(q_H^{FB})$. Furthermore, the profit in this symmetric equilibrium is what the isps could achieve through off-net cost pricing and should give each isp a profit smaller than the upper bound. This ends the proof.

Hence, the upper bound of the joint profits characterized in Proposition 5 can be always achieved by neutral networks.

6.3 Comparison of quality choices

Figure 2 shows the optimal quality schedules for both network regimes.

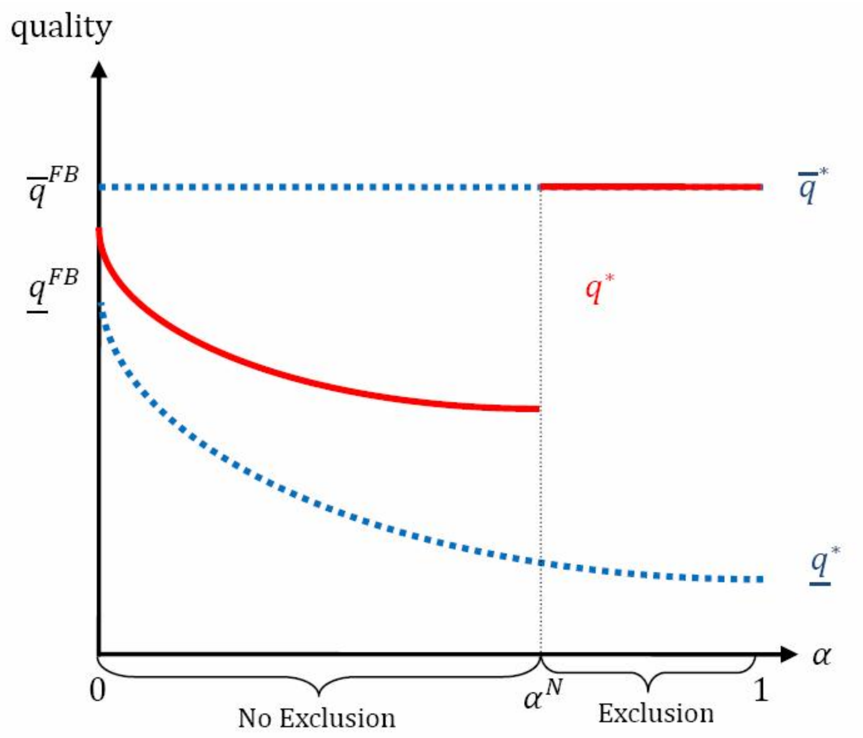

Figure 2: The quality choice

6.4 Bill and Keep

We here discuss access charges chosen by ISPs. In particular, we study the condition under which ISPs choose bill and keep (i.e. zero access charge), which is the current norm. In reality, implementing a negative access charges is difficult since then by sending artificially generated traffics to a rival ISP, an ISP can make positive access revenue. For this reason, we classify the case that a negative access charge is required into bill and keep.

Let us first study the access charges in the first-best world. The prices described in Proposition 1 are the unique ones that implement the first-best outcome under the budget constraint: each ISP and each CP realizes zero profit and the marginal consumer is indifferent between subscribing and not subscribing. In particular, the price charged for each CP of type θ to use service of quality $q^{FB}(\theta)$ is $\alpha\theta u(q^{FB}(\theta))$. Suppose now that a social planner chooses access charges to implement these prices for CPs through off-net cost pricing. Then, the socially optimal access charge is given by

$$a^{FB}(\theta) = \alpha\theta u(q^{FB}(\theta))/q^{FB}(\theta) - c_O.$$

Therefore, we have

$$\text{Bill \& Keep (i.e., } a^{FB}(\theta) \leq 0) \quad \text{iff} \quad \alpha\theta u(q^{FB}(\theta))/q^{FB}(\theta) - c_O \leq 0.$$

It shows that bill and keep is optimal if α is low enough. When α is small enough, most surplus from interaction between a consumer and a CP is captured by consumer. Therefore, the social planner finds it optimal to subsidize content side by charging access charges below termination cost.

Propositions 4 and 5 show that this property is qualitatively preserved when access charges are chosen by ISPs who maximize their joint profits, regardless of whether networks are neutral or not. ISPs have an incentive to subsidize content side while making profits from consumer side when α is small. In the case of non-neutral networks, we have

$$a_H^* < a^{FB}(\theta_H), \qquad a_L^* > a^{FB}(\theta_L).$$

a_H^* is too low because of the rent given to high type CPs whereas a_L^* is too high because of the downward distortion in the quality for low type CPs. This implies that bill and keep is chosen too often for high type CPs while it is chosen too few times for low type CPs. In the case of neutral networks, if we assume $\alpha^N\theta_H u(q_H^{FB})/q_H^{FB} > c_O$ (i.e. conditional on exclusion of low type CPs, bill and keep is never privately optimal), then we have

$$a^* < a^{FB}(\theta_L).$$

In addition, if $\left(\theta_H u(q_H^*) - \theta_\Delta u(q_L^*(\alpha))\right)/q_H^* > \theta_L u(\widetilde{q}(\alpha))/\widetilde{q}(\alpha)$ holds,

$$a^* < a_H^*.$$

Hence neutral networks charge bill and keep more often than non-neutral networks.

7 Social Welfare: Neutral vs. Non-Neutral Networks

Recall that a neutral network cannot outperform a non-neutral network if either $\alpha = 1$ or $\alpha = 0$, that is, for the two extreme cases of full or zero surplus extraction. For $\alpha = 1$, the neutral network without second-degree price discrimination results in the exclusion of low type CPs, while the non-neutral network does not entail such exclusion. For $\alpha = 0$, the neutral network provides a suboptimal quality for both types of CPs, while the non-neutral network provides the first-best quality for each type of CP. Does this mean that a neutral network is always dominated by a non-neutral one? Interestingly, we find that the neutral network can result in a higher social welfare when we turn our eyes to some intermediate $\alpha \in (0, 1)$.

The intuition is not esoteric. Basically, a non-neutral network generates higher profits from content side compared to a neutral network. This means that the non-neutral network is inclined to distort the quality for the low type CP to extract more surplus compared to the neutral network. The incentive to such a distortion increases in the content-side surplus, α. Thus, the neutral network with $\alpha = 1$ can yield a higher social welfare than the corresponding non-neutral network. However, under the neutral network there will be a critical level of α^N such that the ISPs entail the exclusion of low-type CPs for $\alpha > \alpha^N$. If the neutral network entails no exclusion and the quality distortion effect is high enough, the social welfare may be higher in the neutral network.

Let us show our intuition in a rigorous fashion. The social welfare in the non-neutral network with optimal quality choices is given by

$$W^* = N^* \cdot \omega^* - T(N^*) \tag{18}$$

where

$$\omega^* = u_0 + \sum_{k=H,L} \nu_k[\theta_k u(q_k^*) - cq_k^*] = u_0 + \nu(\theta_H u(q_H^{FB}) - cq_H^{FB}) + (1-\nu)(\theta_L u(q_L^{SB}(\alpha)) - cq_L^{SB}(\alpha))$$

denotes the net social surplus per consumer at the optimal quality choices. Recall that the number of consumers is given by $N^* = 2\left[\frac{1}{2} + \lambda\frac{(1+2H)\cdot\widehat{\Pi}(\alpha)-t}{1+4H}\right]$ where $\widehat{\Pi}(\alpha) = \omega^* - \nu\alpha\theta_\Delta u(q_L^{SB}(\alpha))$. We derive the first-order derivative of the social welfare with respect to α as

$$\begin{aligned}\frac{dW^*}{d\alpha} &= N^* \cdot (1-\nu)(\theta_L u'(q_L^{SB}) - c)\frac{dq_L^{SB}}{d\alpha} \\ &\quad + \left(\omega^* - T'(N^*)\right)\cdot 2\lambda\frac{(1+2H)}{1+4H}\left\{\left[(1-\nu)(\theta_L u'(q_L^{SB}) - c) - \nu\alpha\theta_\Delta u'(q_L^{SB})\right]\frac{dq_L^{SB}}{d\alpha} - \nu\theta_\Delta u(q_L^{SB})\right\}\end{aligned}$$

Using the first-order optimal quality condition for the low-type CP, $(1-\nu)(\theta_L u'(q_L^{SB}) - c) - \nu\alpha\theta_\Delta u'(q_L^{SB}) = 0$, we obtain more simplified expression as:

$$\frac{dW^*}{d\alpha} = N^* \cdot \nu\alpha\theta_\Delta u'(q_L^{SB})\frac{dq_L^{SB}}{d\alpha} - (\omega^* - T'(N^*))2\lambda\frac{(1+2H)}{1+4H}\nu\theta_\Delta u(q_L^{SB}). \tag{19}$$

As we derived in (7), the quality distortion increases in α, i.e., $\frac{dq_L^{SB}}{d\alpha} < 0$: the first term of (19) takes on a negative value. We can know $\omega^* - T^{*\prime} \geq 0$ because there are fewer subscribing consumers with any quality distortion; the equality holds only when the first-best quality for each type of CPs is provided as in (4). So, the second term of (19) is also negative including its leading negative sign. Thus, we find that the social welfare in the non-neutral network decreases in the degree of CPs' rent extraction, that is, $\frac{dW^*}{d\alpha} < 0$. The welfare in the non-neutral network starts at the first-best welfare level when $\alpha = 0$. Similarly, the social welfare in the neutral network also decreases in α for any $\alpha < \alpha^N$:

$$\frac{d\widetilde{W}^*}{d\alpha} = \widetilde{N}^* \cdot \alpha\nu\theta_\Delta u'(\widetilde{q}^*)\frac{d\widetilde{q}^*}{d\alpha} - \left(\widetilde{\omega}^* - \widetilde{T}^{*\prime}\right)2\lambda\frac{(1+2H)}{1+4H}\nu\theta_\Delta u(\widetilde{q}^*) < 0. \tag{20}$$

Recall that the quality adjustment with respect to the change in α can be derived as (7) for the non-neutral network and (9) for the neutral one. For the sake of comparison, let us consider any utility function with Arrow-Pratt constant absolute risk aversion (CARA), e.g., $u(q) = A - \frac{B}{r}\exp(-rq)$ where r measures the degree of risk aversion. Then, we can obtain a clear comparison of

$$\left|\frac{dq_L^{SB}}{d\alpha}\right| > \left|\frac{dq^*}{d\alpha}\right| \quad \text{for } \forall\nu \in (0,1) \tag{21}$$

from $-\frac{u'(\widetilde{q}^*)}{u''(\widetilde{q}^*)} = -\frac{u'(q_L^{SB})}{u''(q_L^{SB})} = r$. This implies that the ISPs decrease their quality to the low-type q_L^{SB} more quickly in the non-neutral networks than the uniform quality q^* in the neutral network when the content-side surplus increases. In addition, we find $u'(q_L^{SB}) > u'(\widetilde{q}^*)$ from $q_L^{SB} < \widetilde{q}^*$ for any utility function with $u'' < 0$. Since the non-neutral network provides total surplus at least as high as that under the neutral network, we have $N^* \geq \widetilde{N}^*$ where the equality holds for $\lambda = 0$. Hence, we find that the social welfare more quickly decreases as α increases in the non-neutral networks compared to the neutral network, i.e., $\left|\frac{dW^*}{d\alpha}\right| > \left|\frac{d\widetilde{W}^*}{d\alpha}\right|$ if the market expansion is highly limited ($\lambda \approx 0$).

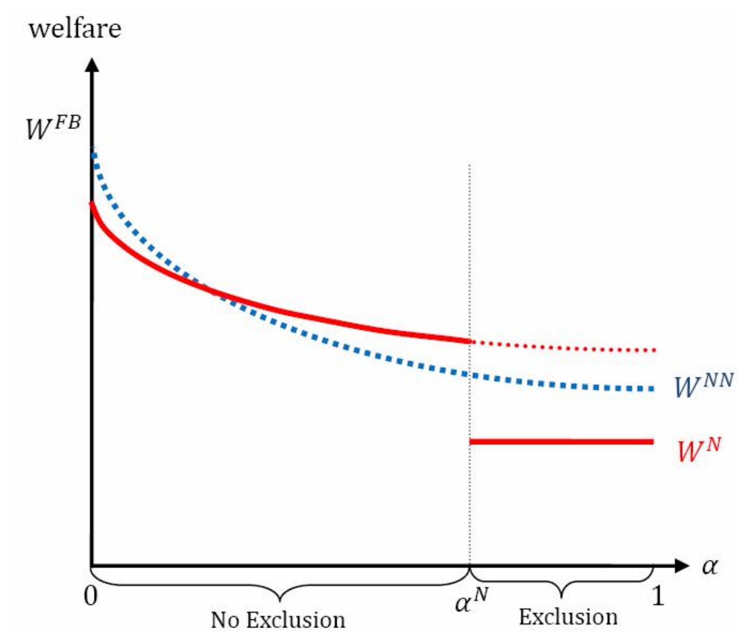

Figure 3: The social welfare

Given this understanding, let us finally compare the level of social welfare under two different network regimes. Recalling $(\theta_L + (1-\alpha^N)\nu\theta_\Delta)u(\widetilde{q}^*) - c\widetilde{q}^* = \nu \cdot (\theta_H u(q_H^{FB}) - cq_H^{FB})$, the per consumer net surplus in the neutral network can be expressed as

$$\begin{aligned} \widetilde{\omega}^* &= (\theta_L + (1-\alpha^N)\nu\theta_\Delta)u(\widetilde{q}^*) - c\widetilde{q}^* + \alpha^N \nu\theta_\Delta u(\widetilde{q}^*) \\ &= \nu \cdot (\theta_H u(q_H^{FB}) - cq_H^{FB}) + \alpha^N \nu\theta_\Delta u(\widetilde{q}^*) \end{aligned}$$

This simplifies the comparison between $\widetilde{\omega}^*$ and ω^* as the comparison between $\alpha^N \nu\theta_\Delta u(\widetilde{q}^*)$ and $(1-\nu)(\theta_L u(q_L^{SB}) - cq_L^{SB})$:

$$\widetilde{\omega}^* > \omega^* \Leftrightarrow \alpha^N \nu\theta_\Delta u(\widetilde{q}^*) > (1-\nu)(\theta_L u(q_L^{SB}) - cq_L^{SB}), \tag{22}$$

Under Assumptions 1-2, per consumer social welfare can be higher in a neutral network relative to a non-neutral network as long as (22) is satisfied. Note that this does not ensure that a neutral network always dominates a non-neutral network at α^N because of $N^* \geq \widetilde{N}^*$. However, we can state that

Proposition 6 *Consider the CARA utility function for which Assumptions 1 and 2, and (22) are satisfied. For a sufficiently small λ, there exists always intermediate level of α, which is smaller than α^N but is greater than zero, under which the total social surplus also can be higher in the neutral network.*

This result suggests that the merit of net neutrality regulation may depend crucially on content providers (CPs)' business models.

8 Concluding Remarks

This paper analyzed competition between interconnected networks when content is heterogeneous in terms of its sensitivity to delivery quality. The heterogeneity of content calls for the multi-tiered Internet to reflect the need for differential treatment of packets depending on its types. With interconnected networks, however, the assurance of a certain level of delivery quality requires cooperation among networks. To address this issue, we have developed a framework of two-sided markets in which ISPs compete with each other to serve as platforms that connect CPs and end consumers. We have considered two regimes under which packets can be delivered: a neutral regime in which all packets are required to be delivered with the same quality (speed) and a non-neutral regime under which ISPs are allowed to offer multi-tiered services with different delivery quality levels. We derived conditions under which social welfare can be higher in a neutral network. The conditions highlight the importance of CPs' business models in the evaluation of net neutrality regulation.

Looking forward, this paper is a first step towards incorporating heterogeneous content in the analysis of interconnection issues. There are many worthwhile extensions that calls for further analysis. One limitation of our analysis is its static nature. We have not analyzed dynamic investment incentives facing ISPs and CPs by assuming away capacity constraint for ISPs and a fixed measure of CPs. The effects of net neutrality regulation on ISPs' capacity expansions and CPs' entry decisions would be an important issue.[8]

We have also assumed two symmetric networks. However, mobile networks are becoming an

[8] Choi and Kim (2010) addresses the dynamic investment issue. However, they consider a monopolistic ISP, and thus the issues of interconnection and access charge regulation do not arise.

increasingly important channel for Internet content delivery. Fixed and mobile Internet networks are inherently different in many dimensions, most importantly the scarcity of bandwidth for mobile networks imposed by physical laws. These differences are also recognized by the recent FCC rule on net neutrality. The new FCC rule, announced on December 21, 2010, reaffirmed its commitment to the basic principle of net neutrality by prohibiting ISPs from "unreasonable discrimination" of web sites or applications, but wireless telecommunications were exempted from such anti-discrimination rules. It would be an important research agenda to reflect such differences in networks and analyze the issue of interconnections between asymmetric networks. In particular, our model may lend a new justification for asymmetric regulation between fixed and mobile networks. For instance, imagine a situation in which the mobile networks are more constrained in their capacity and expansion possibilities. The network operators thus may prefer to serve only the high type CPs under neutral network, instead of providing somewhat jittery content delivery by serving all types. If CPs in the mobile networks adopt business models of more content-usage based charge system that enables them to extract more surplus from consumers than the ad-financed system, our model suggests that net neutrality regulation may be beneficial for fixed networks but not for mobile networks.

Finally, we assumed a homogeneous and exogenous business model by assuming the same level of surplus extraction (parametrized by α) for CPs. The analysis can be extended to heterogeneous business models or business models can be endogenously derived.

References

[1] Armstrong, M. "Network Interconnection in Telecommunications." Economic Journal, Vol. 108 (448), 1998: pp 545-564.

[2] Armstrong, M. "Competition in Two-Sided Markets," Rand Journal of Economics, vol. 37(3), Autumn 2006, pp. 668-691.

[3] Armstrong, M. and J. Wright. "Mobile Call Termination." Economic Journal, Vol. 119(538), 2009, F270-F307.

[4] Atkinson, J and C. Barnekov, "A Competitively Neutral Approach to Network Interconnection," OPP Working Paper No. 34 (2000).

[5] Caillaud, B. and B. Jullien. "Chicken and Egg: Competition among Intermediation Service Providers." Rand Journal of Economics, Vol.34(2), Summer 2003, pp. 309-328.

[6] Cheng, H.K., S. Bandyopadhyay, and H. Guo. "The Debate on Net Neutrality: A Policy Perspective." Information Systems Research, 22(1), March 2011, pp. 60-82.

[7] Choi, J.P. and B-C. Kim. “Net Neutrality and Investment Incentives.” Rand Journal of Economics, Vol. 41(3), 2010, pp. 446-471.

[8] Crémer, J., P. Rey, and J. Tirole. “Connectivity in the Commercial Internet.” Journal of Industrial Economics, Vol.48(4), 2000, pp. 433-472.

[9] Dessein, W. “Network Competition in Nonlinear Pricing.” Rand Journal of Economics, Vol.34(4), Winter 2003, pp. 593-611.

[10] Economides, N. and B. Hermalin. “The Economics of Network Neutrality.” (2010). Rand Journal of Economics, forthcoming.

[11] Economides, N. and J. Tåg, “Net Neutrality on the Internet: A Two-sided Market Analysis” NET Institute Working Paper No. 07-14, NYU Law and Economics Research Paper

[12] Gans, J.S., and S.P. King. “Using ‘Bill and Keep’ Interconnect Arrangements to Soften Network Competition.” Economics Letters, Vol. 71, 2001, pp. 413-420.

[13] Hagiu, A. and R.S. Lee, ”Exclusivity and Control,” Journal of Economics and Management Strategy, 2011, pp. 679-701.

[14] Hahn, J.H. “Network Competition and Interconnection with Heterogeneous Subscribers.” International Journal of Industrial Organization, Vol. 22(5), 2004, pp. 611-631.

[15] Hermalin, B. and M. Katz. “Network Interconnection with Two-Sided User Benefits.” Mimeo, UC Berkeley, (2001).

[16] Hermalin, B. and M. Katz. “Sender or Receiver:Who Should Pay to Exchange an Electronic Message?” Rand Journal of Economics, Vol. 35(3), 2004, pp. 423-448.

[17] Hermalin, B. and M. Katz, “The Economics of Product-Line Restrictions With an Application to the Network Neutrality Debate.” Information Economics and Policy, Vol.19(2), 2007, pp.215-248.

[18] Hurkens, S. and D.S. Jeon, 2011. ”Promoting Network Competition by Regulating Termination Charges.” Mimeo, Toulouse School of Economics.

[19] Jeon, D.S. and S. Hurkens. “A Retail Benchmarking Approach to Efficient Two-Way Access Pricing: No Termination-based Price Discrimination.” Rand Journal of Economics, Vol. 39(3), Autumn 2008, pp. 822-849.

[20] Jeon, D.S., J.J. Laffont and J. Tirole. “On the ‘Receiver-Pays’ Principle.” Rand Journal of Economics, Vol.35(1) Spring 2004, pp. 85-110.

[21] Jullien, Bruno. ”Participation Constraints in Adverse Selection Models”, Journal of Economic Theory, vol. 93, n. 1, 2000, p. 1-47.

[22] Kocsis, V. and P. de Bijl, “Network Neutrality and the Nature of Competition between Network Operators.” International Economics and Economic Policy, vol. 4, 2007, Special Issue on the Digital Economy and Regulatory Issues, pp. 159-184.

[23] Laffont, J.J., S. Marcus, P. Rey and J. Tirole. “Internet Interconnection and the Off-Net-Cost Pricing Principle.” Rand Journal of Economics, Vol. 34, 2003, pp. 370-390.

[24] Laffont, J.J, S. Marcus, P. Rey and J. Tirole, "Internet Peering." American Economic Review, vol. 91, n. 2, May 2001, p. 287-291.

[25] Laffont, J.J., P. Rey and J. Tirole. "Network Competition I: Overview and Nondiscriminatory Pricing." Rand Journal of Economics, Vol. 29(1), 1998a, pp. 1-37.

[26] Laffont, J.J., P. Rey and J. Tirole. "Network Competition II: Price Discrimination." Rand Journal of Economics, Vol. 29(1), 1998b, pp. 38-56.

[27] Maskin, E., and J. Riley, "Monopoly with Incomplete Information." Rand Journal of Economics, Vol. 15, 1984, pp. 171-196.

[28] Mussa, M., and S. Rosen. "Monopoly and Product Quality." Journal of Economic Theory, Vol. 18, 1978, pp. 301-317.

[29] Njoroge, P., A.E. Ozdaglar, N.E. Stier-Moses, , G.Y. Weintraub. "Investment in Two Sided Markets and the Net Neutrality Debate." Columbia Business School DRO (Decision, Risk and Operations) Working Paper No. 2010-05.

[30] Schuett, F. "Network Neutrality: A Survey of the Economic Literature." Review of Network Economics: Vol. 9(2), Article 1.

[31] Valletti, T. and C. Cambini. "Investments and Network Competition." Rand Journal of Economics, Vol. 36(2), Summer 2005, pp. 446-467.

[32] Wright, J. "Access Pricing under Competition: An Application to Cellular Networks." Journal of Industrial Economics, Vol. 50, 2002, pp. 289-315.

[33] Wu, T. "Network Neutrality, Broadband Discrimination." Journal on Telecommunications and High Technology Law 2, 2003a, pp.141-178.