Net Neutrality, Business Models, and Internet Interconnection[*]

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

July 18, 2014

Abstract

We analyze the effect of net neutrality regulation in a two-sided market framework when content is heterogeneous in its sensitivity to delivery quality. We characterize the equilibrium in a neutral network constrained to offer the same quality vis-à-vis a non-neutral network where Internet service providers (ISPs) are allowed to engage in second degree price discrimination with a menu of quality-price pairs. We find that the merit of net neutrality regulation depends crucially on content providers' business models. More generally, our analysis can be considered a contribution to the literature on second-degree price discrimination in two-sided platform markets.

JEL Codes: D4, L1, L5

Key Words: Net neutrality, Two-sided markets, Second-degree price discrimination, Content providers' business models, Internet interconnection

[†]School of Economics, University of New South Wales, Sydney, NSW 2052, Australia, Department of Economics, Michigan State University and School of Economics, Yonsei University. E-mail: choijay@gmail.com.

[‡]Toulouse School of Economics (GREMAQ, IDEI), Manufacture de Tabacs, 21 Allée de Brienne, 31000, Toulouse, France and School of Economics, Sungkyunkwan University, 53 Myungryun-Dong 3 Ga, Jongro-Gu, Seoul, Korea, 110-745. 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

With the emergence of the Internet as a fundamental infrastructure for communication, information and commercial activities, "net neutrality" has become one of the most important regulatory policy issues. In essence, net neutrality is a principle that Internet service providers (ISPs) must treat 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.[1] From its inception, the Internet has implicitly been governed by this principle of equal access to all types of content.

However, with the emergence of various online multimedia services that demand a significant amount of network bandwidth, network congestion and efficient management of network resources have become important policy issues. In particular, content and applications differ in their sensitivity to delay in delivery. For instance, data applications such as e-mail can be relatively insensitive to moderate delivery delays from users' viewpoints. By contrast, streaming video/audio or Voice over Internet Protocol (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 is not an efficient way to utilize the network. In addition, with the reclassification of Internet transmissions from the category of "telecommunications services" to the category of "information services" by the Federal Communications Commission (FCC) in 2005, ISPs are no longer considered as a common carrier utility and subject to nondiscrimination restrictions. As a result, major ISPs such as AT&T, Comcast and Verizon have expressed an interest in offering multi-tiered services that would provide content providers paying a premium with a faster access to their networks. In response, the Federal Communications Commission (FCC) implemented three basic "Open Internet Rules" in 2010 to maintain the status quo of net neutrality regime: transparency, no blocking, and no unreasonable discrimination.[2] However, on January 14, 2014, the U.S. Court of Appeals for the D.C. Circuit vacated the non-discrimination and anti-blocking rules of the FCC's Open Internet Order based

---

[1]There is no universally accepted definition of net neutrality. The definition given here is similar to what Krämer, Wiewiorra and Weinhardt (2013) calls "strict net neutrality" and consumer groups advocate. For various definitions of net neutrality and more details about the debate on net neutrality, see the survey paper by Krämer, Wiewiorra, and Weinhardt (2013).

[2]According to the FCC website (http://www.fcc.gov/guides/open-internet), "The 'Open Internet' is the Internet as we know it... Network, or 'net,' neutrality is just another way of referring to Open Internet principles."

on a technical reason that the FCC does not have the authority to regulate how ISPs grant access to content.[3]

In this net neutrality debate, consumer groups and policy makers are mostly concerned with the possibility that ISPs may "cut deals with some content providers for faster access while forcing others into the slow lane" (*USA TODAY*, January 16, 2014) by offering multi-tiered services, whereas ISPs argue that they should be able to manage their networks as they see fit to justify their investments in upgrading their networks. To address this issue, we analyze the effect of net neutrality regulation in a two-sided market framework in which Internet service providers (ISPs) serve as platforms that connect content providers (CPs) and end consumers. To reflect the current debate, we focus on price discrimination based on the speed of access as a particular deviation from the net neutrality principle.

On the CP side, there is a continuum of heterogeneous content/application providers. CPs' content differs in its sensitivity to delivery quality: for a clear exposition, we consider two types of CPs. We compare the equilibrium in a neutral network constrained to offer the same quality with the one in a non-neutral network where ISPs are allowed to engage in second degree price discrimination with a menu of quality-price pairs. While the difference in the congestion sensitivity of content providers justifies the need to provide multiple lanes for different delivery qualities, we find that a neutral treatment can be welfare-enhancing depending on the CPs' relative share of total surplus.

As the degree of consumer surplus that CPs extract is primarily affected by the business models that CPs use, our results reveal the importance of CPs' business models in assessing the effects of net neutrality regulation. More generally, our findings contribute to the literature on second degree price discrimination in two-sided markets. We show that how a platform's second-degree price discrimination fares against no discrimination depends on the relative allocation of each group's surplus in a two-sided market.[4]

---

[3]To quote Circuit Judge David Tatel, "Given that the commission has chosen to classify broadband providers in a manner that exempts them from treatment as common carriers, the Communications Act expressly prohibits the commission from nonetheless regulating them as such. Because the commission has failed to establish that the anti-discrimination and anti-blocking rules do not impose per se common carrier obligations, we vacate those portions of the *Open Internet Order*." See *Verizon v. FCC & USA*, No. 11-1355 (D.C. Cir.). As of this writing, the FCC has proposed revised rules that would prevent ISPs from deliberately blocking or slowing content delivery, but allow CPs to pay for a faster lane of service (FCC 14-61, May 15, 2014). Netflix has already cut a deal with Comcast for a faster streaming service.

[4]While our model is primarily motivated by the net neutrality debate, its implications may

To establish our main intuition, we start with a monopolistic ISP facing homogeneous consumers as a basic model. In particular, we first consider a scenario in which the surpluses from interactions between the CPs and end consumers can be entirely appropriated by one-side of the market. In such a scenario, price discrimination is socially more efficient than neutral treatment; thereby we stack the deck against the neutral regime. Nonetheless, we show that the social welfare can be higher with neutrality regulation when surplus extraction is neither full nor zero.

The intuition for this main result is as follows. When choosing the quality for low type CPs, the ISP faces a trade-off arising from the two-sided nature of its business. A downward distortion in the quality for low type CPs has a *benefit* of extracting more rent of high type CPs, on the CP side, and a *cost* of reducing the consumer surplus that the ISP can extract, on the consumer side. This trade-off implies that as the ISP focuses on extracting consumer surplus rather than CPs' surplus, there will be less distortion in quality. Even if the ISP can extract full consumer surplus regardless of the neutrality regulation in place, the ISP tends to focus more on extracting CPs' surplus in the non-neutral network than in the neutral network since the ISP has more instruments to extract CPs' surplus in a non-neutral network than in a neutral network. This is why welfare can be higher in the neutral network than in the non-neutral network.

More specifically, consider first the extreme case in which consumers take the entire surplus generated by content delivery and CPs' share is zero. Then, the ISP will provide the *first best* quality for each type of CPs in a non-neutral network because this allows them to extract the highest consumer surplus from subscription fees. By contrast, a suboptimal single quality is provided in a neutral network due to the regulatory restriction. Thus, in this case the non-neutral network yields a strictly higher social welfare than the neutral network. As the CPs' relative share of total surplus increases, extracting high type CPs' rent becomes more and more important. As a consequence, the social welfare ranking between the two regimes would be reversed due to the accelerated quality distortion against low type CPs in the non-neutral network, provided that the neutral network still serves both types of CPs. As the CPs' relative share further increases, the exclusion of low type CPs may

---

extend to any market environments in which a two-sided platform has a substantial market power in one side of the market and practice menu pricing in the other side. One example is Korean mobile operators which used subsidies to handsets as a price discrimination instrument against mobile phone customers. The major wireless carriers completely control the distribution of mobile phones. The Korea Communications Commission (KCC) decided to regulate on smartphone subsidies to limit price discrimination by capping the amount of subsidies per handset to KRW 270,000.

occur under a single quality provision and the non-neutral network reclaims a higher social welfare. This is because the non-neutral network still serves low type CPs while high type CPs are offered the first best quality in both network regimes. This shows that the welfare comparison between the two different network regimes may reveal a non-monotonic relationship with respect to the relative allocation of total surplus between CPs and end consumers.

Can this result still be meaningful even in the real Internet environment in which multiple ISPs intensely compete and consumers are heterogeneous in their preferences toward different ISPs? Our answer to this inquiry is positive.

With competing ISPs and heterogeneous consumers, several additional issues arise. In particular, 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. In addition, 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. Guaranteeing a specified quality (speed) of content delivery requires cooperation from other networks when content providers and end users belong to different networks. We assume that the ISPs agree on the delivery quality and reciprocal access charge(s) for the delivery of other ISPs' traffic that terminate on their own networks. We assume that CPs can multi-home whereas consumers single-home and constitute competitive bottlenecks.[5] However, because of the interconnection arrangement, a CP can deliver content to consumers subscribed to different ISPs without subscribing to multiple ISPs.

We find that any equilibrium with interconnection is governed by the so-called "off-net cost pricing principle" on the CP side. The off-net cost pricing principle was discovered by Laffont, Marcus, Rey, and Tirole (hereafter LMRT, 2003) and Jeon, Laffont and Tirole (2004). It means that network operators set prices for their customers as if their customers' traffic were entirely off-net. We find a novel equivalence result : competing ISPs agree on access charges and delivery qualities to maximize the same objective as a monopolistic ISP facing homogeneous consumers, even if this implies intensified competition on the consumer side. By using this equivalence, we show that the qualitative results derived with a monopolistic ISP naturally extend to the case of competing ISPs with interconnection.[6]

---

[5] See Armstrong (2006) for various modes of competition in two-sided markets.

[6] Our analysis of interconnection can be of independent interest and contributes to the interconnection literature by considering heterogeneous content and price discrimination. We generalize

4

The remainder of the paper is organized as follows. In Section 2, we discuss related literature and our contribution. In Section 3, we set up a basic model of two-sided markets with a monopolistic ISP and homogeneous consumers. We analyze the effects of price discrimination across different types of content with a menu of contracts. In Section 4, we compare a neutral network with a non-neutral network in terms of quality choices and social welfare and derive conditions under which the neutral regime can provide higher welfare than the non-neutral one. Section 5 extends the analysis to competing interconnected ISPs facing heterogeneous consumers. We derive a central equivalence result between competing ISPs and a monopolistic ISP. This equivalence result ensures the robustness of our result to the introduction of ISP competition and interconnection. Section 6 contains our concluding remarks.

2 Related Literature

Our research 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.[7]

Hermalin and Katz (2007) examine a situation in which ISPs serve as platforms to connect CPs with consumers. Without any restrictions, ISPs can potentially offer a continuum of vertically differentiated services to a continuum of types of CPs. With net neutrality regulation, ISPs are required to provide a single tier of Internet service. They compare the single service equilibrium with the multi-service equilibrium. One novelty of our paper with respect to Hermalin and Katz (2007) is that we analyze how the relative merit of allowing second-degree price discrimination depends on CPs' business models that determine the relative allocation of total surplus between CPs and end consumers. In addition, we extend the analysis to allow for interconnection between competing ISPs.[8]

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*

---

the finding of LMRT to a setting of heterogeneous content with different delivery qualities across content.

[7]See Lee and Wu (2009), Schuett (2010), Krämer, Wiewiorra, and Weinhardt (2013) for the surveys about economics literature on network neutrality.

[8]Although they model a two-sided market, they would obtain the same qualitative results in a one-sided market. In contrast, in our model, a non-neutral network always generates higher welfare than a neutral network if the market is one-sided while this can be reversed if the market is two-sided.

ISPs. Choi and Kim (2010) analyze the effects of net neutrality regulation on investment incentives of a monopoly ISP and CPs. They show that ISPs may invest less in capacity in a non-neutral network than in a neutral network because expanding capacity reduces the CPs' willingness to pay for having a prioritized service. 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. They show that the ability to price discriminate enhances incentives to invest, creating a trade-off between static and dynamic efficiencies. As these papers consider a monopolistic ISP, the interconnection and competition issues do not arise.

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 and assume that a CP has access only to end users connected to the same ISP. Economides and Tåg (2012) also consider both a monopolistic ISP and duopolistic ISPs. Once again the issue of Internet interconnection is not considered as they focus on how net neutrality regulation as a zero pricing rule affects pricing schemes on both sides of the market and social welfare.

Our research also relates to 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 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 types. In this setting, we analyze how the quality levels and access charges are determined (depending on CPs' business models and on whether there exists net neutrality regulation) and find a novel equivalence result between competing ISPs and a monopolistic ISP.

There is a large literature on interconnection in the telecommunication market, initiated by Armstrong (1998) and Laffont, Rey, and Tirole (1998a,b). These researchers show that if firms compete in linear prices, they agree to set interconnection charges above associated costs to obtain the joint profit-maximizing outcome and derive the welfare-maximizing interconnection charge that is lower than the privately negotiated level. They also show that the nature of competition can be altered significantly depending on whether or not two-part tariffs or termination-based price discrimination are employed as price instruments. Their models, however, are devoid of the

issue of transmission of quality because all calls are homogeneous. In contrast, we consider heterogeneous types of CPs requiring different transmission qualities and analyze quality distortions associated with net neutrality regulations. Armstrong (1998) and Laffont, Rey, and Tirole (1998a,b) consider inelastic subscription of consumers while we consider elastic subscription demand. If we consider inelastic subscription, we find a profit neutrality result.[9] Although some papers study elastic subscription demand in the literature on interconnection in the telecommunications market (see Armstrong and Wright, 2009, Dessein, 2003, Hurkens and Jeon, 2012), to our knowledge, our result of equivalence between competing ISPs and a monopolistic ISP is new.

3 A Monopolistic ISP in a Two-sided Market

3.1 ISPs, CPs, and Consumers

We consider a model of a two-sided market to analyze the effects of price discrimination on various market participants and social welfare. To be concrete, we consider a monopolistic ISP that serves as a platform 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 take up small 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.[10]

There is a continuum of CPs whose mass is normalized to one. We consider a simple case of CP heterogeneity. There are two types of CPs: $\theta \in \{H, L\}$, with $\Delta = H - L > 0$. The measure of θ type CP is denoted by ν_θ, where $\nu_H = \nu$ 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

---

[9] More precisely, in an earlier version of this paper where we assume that consumers are distributed over a Hotelling line, we showed that each ISP's profit is constant and equal to the Hotelling profit regardless of quality levels and access charges.

[10] For instance, e-mail exchanges between consumers take up trivial volumes. With the developments of VoIP, file sharing, and online gaming services, the absolute volume of consumer-to-consumer traffic is increasing, but still expected to constitute a relative decrease of the percentage of total traffic due to explosive growth in Internet video streaming and downloads. See Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2012-2017, available at http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360.pdf

context, quality means speed and reliability of content delivery. Let q_θ denote the quality of delivery associated with content of type θ. 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 $u'' < 0$ with the Inada condition $\lim_{q \to 0} u'(q) = \infty$. According to our utility formulation, θ reflects the sensitivity of content to delay, with higher valuation content being more time/congestion sensitive.[11] 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 the CPs' business model. We have in mind two sources of revenue for CPs: micropayments and advertising revenue.[12] For instance, the parameter α would be higher if CPs can extract surplus from consumers via micropayments in addition to advertising revenues. If the CPs' revenue source is limited to advertising, α can be relatively low. We later show that the CPs' business model, captured by α, plays an important role in assessing the effects of net neutrality regulations.

A monopolistic ISP provides content delivery service 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 $q \geq 0$.[13] We consider two different regimes under

---

[11]For our analysis, the type space is essentially about the sensitivity to transmission quality; perfect correlation between the value of content and its sensitivity to transmission quality is not necessary. To illustrate this, consider an alternative formulation that explicitly accounts for the transmission delay costs. For instance, assume that all CPs provide the same gross utility of v (i.e., there is only one type in terms of the content value), but this value is reduced with a delay. Consider an additive specification in which the total surplus generated from interaction between a consumer and a CP of type θ is equal to $v - \theta d$, where d is delay time. With this additive specification (and independence between the content value and the sensitivity type parameter), we can derive exactly the same qualitative results as in the paper.

[12]Our simplifying assumption is that CPs are homogeneous in all dimensions except for their type θ. This implies that they use the same business model; otherwise, CPs have two-dimensional types (θ and business model) from the ISP's point of view. In an ad-based business model, we can assume that the advertising revenue is proportional to consumer gross utility, which in turn is proportional to $\theta u(q)$. Then, the total surplus is given by $b\theta u(q)$ where b is a positive constant. Hence, by redefining $b\theta$ as θ', we are back to our original formulation. See Appendix B for a micro-foundation of such an advertising model.

[13]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 non-linear. 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, the utility function with the normalized cost still remains concave.

which ISPs can deliver content: a neutral regime or a non-neutral regime. Under a non-neutral regime, the ISP can offer multiple classes of services that differ in delivery quality. We assume that the ISP is unable to practice first-degree price discrimination across content providers depending on content types, but 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 for low-type CPs. In a neutral regime or in the presence of net neutrality regulation, ISP i is constrained to offer a single uniform delivery quality q.

To focus on the effects of the ISP's price discrimination against CPs, we assume homogeneous consumers. This setup allows the ISP to extract the whole consumer surplus. However, we extend the analysis later to allow for competing interconnected ISPs facing heterogeneous consumers with elastic participation, and establish an equivalence result between the monopolistic outcome of the basic model and the outcome of competing ISPs in terms of offers made to the content side.

Let $U(\alpha)$ denote the gross utility a consumer derives from the content side by subscribing to the ISP. We have $U(\alpha) = \underline{u} + (1 - \alpha)\sum_{\theta}\nu_{\theta}\theta u(q_{\theta})$, where $\underline{u}$ is the intrinsic utility associated with the Internet connection. A consumer's net utility from subscribing to the ISP, u, is given by

$$u(\alpha, f) = U(\alpha) - f, \tag{1}$$

where f is the subscription price charged by the ISP. We normalize, without loss of generality, the total measure of consumers to one. Finally, we assume that the monopoly ISP simultaneously announces the price-quality pairs for CPs and the fee for consumers.[14]

Before analyzing market outcomes under various regimes, we first analyze the first-best outcome as a benchmark. It is clear that the socially optimal quality should maximize $\theta u(q) - cq$ and hence, the first-best quality level for CPs of type θ, denoted q_{θ}^{FB}, is determined by the following condition:

$$\theta u'(q_{\theta}^{FB}) = c. \tag{2}$$

---

[14]The optimal outcome chosen by the monopoly ISP in this simultaneous pricing is the same as the one chosen in a sequential pricing in which it first chooses the price-quality pairs for CPs and then the fee for consumers.

9

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. 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.

Our modeling strategy is driven by institutional features of the Internet market. First, in many countries including the US, customers predominantly pay flat fees for their access to the Internet regardless of the traffic amount they generate.[15] In addition, even if consumers pay different prices depending on the amount of data they stream, differential pricing on the consumer side seems to be non-controversial and even considered necessary to manage over-loaded Internet traffic. Thus, we focus on two-tiered services with price discrimination in only the content provider side.[16]

3.2 Non-neutral Network and Second-degree Price Discrimination

Let $\{(p_H(\alpha), q_H(\alpha)), (p_L(\alpha), q_L(\alpha))\}$ be the menu of contracts offered to CPs which satisfies the incentive and participation constraints of CPs (defined below). Then, each consumer's gross utility is given by $U(\alpha) = \underline{u} + (1-\alpha)\sum_{\theta}\nu_\theta\theta u(q_\theta)$, which can be fully extracted by a subscription fee f as consumers are homogeneous. The ISP's profit from the content side is $\pi^{CP} = \sum_{\theta}\nu_\theta[p_\theta - cq_\theta]$. The overall profit for the ISP can be written as $\Pi^M(\alpha) = U(\alpha) + \pi^{CP}$. Thus, the monopolistic ISP's mechanism design problem can be described as:

$$\max_{(p_\theta, q_\theta)} \Pi^M(\alpha) = \underline{u} + \sum_{\theta}\nu_\theta[p_\theta + (1-\alpha)\theta u(q_\theta) - cq_\theta]$$

subject to

$$
\begin{aligned}
IC_H &: \quad \alpha H u(q_H) - p_H \geq \alpha H u(q_L) - p_L; \\
IC_L &: \quad \alpha L u(q_L) - p_L \geq \alpha L u(q_H) - p_H; \\
IR_H &: \quad \alpha H u(q_H) - p_H \geq 0; \\
IR_L &: \quad \alpha L u(q_L) - p_L \geq 0,
\end{aligned}
$$

---

[15]See Krämer, Wiewiorra, and Weinhardt (2013) for the "flat rate trap" the ISPs are in on the consumer side of the market.

[16]However, we verify the robustness of our results in a simple extension of the model with two symmetric sides in which both the end user and the CP sides are heterogeneous and the ISP practices menu pricing on both sides (the details are available upon request).

where IC_θ and IR_θ refer to type θ CPs' incentive compatibility constraint and individual rationality constraint, respectively.

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

$$p_H = \alpha H u(q_H) - \alpha \Delta u(q_L); \qquad p_L = \alpha L u(q_L). \tag{3}$$

This leads to the following reduced problem

$$\max_{\{q_H, q_L\}} \Pi^M(\alpha, q_H(\alpha), q_L(\alpha)) = \underline{u} + \sum_\theta \nu_\theta [\theta u(q_\theta) - c q_\theta] - \alpha \nu \Delta u(q_L).$$

The objective in the reduced program shows that the ISP extracts full surplus except for the rent to high type CPs, which is given by $\alpha \nu \Delta u(q_L)$. Let $\{(p_H^*(\alpha), q_H^*(\alpha)), (p_L^*(\alpha), q_L^*(\alpha))\}$ be the menu of contracts chosen by the ISP under a non-neutral network. From the first order conditions, we find that the optimal quality for the high type is determined by $H u'(q_H^*) = c$ for any α, which is equal to the first-best level, regardless of α, i.e., $q_H^* = q_H^{FB}$. By contrast, the low type CPs' quality is characterized by

$$\left(L - \frac{\nu}{1-\nu} \cdot \alpha \Delta \right) u'(q_L^*(\alpha)) = c. \tag{4}$$

As in the standard mechanism design problem, there is a downward distortion in quality for the low type, that is, $q_L^*(\alpha) \leq q_L^{FB}$ with the equality holding only for $\alpha = 0$.

We assume that if CPs extract all the surplus from consumers (i.e., $\alpha = 1$), the monopoly ISP prefers serving both types under second-degree price discrimination.

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

Under the Inada condition, $q_L^*(\alpha = 1) > 0$ is equivalent to $L > \frac{\nu}{1-\nu}\Delta$. Assumption 1 ensures that $q_L^*(\alpha) > 0$ for any $\alpha \in [0, 1]$ because total differentiation applied to (4) shows that the low-type quality is decreasing in α :

$$\frac{dq_L^*}{d\alpha} = \frac{\nu \Delta u'(q_L^*)}{((1-\nu)L - \alpha\nu\Delta) u''(q_L^*)} < 0. \tag{5}$$

For a given q_L, the rent obtained by a high type CP increases with α. Hence, as the CPs' share of surplus (i.e., α) increases, the ISP has more incentives to distort the quality for the low type. From the envelope theorem, the maximized objective under

a non-neutral network strictly decreases with α:

$$\frac{d\Pi^M(\alpha, q_H^*(\alpha), q_L^*(\alpha))}{d\alpha} = -\nu \cdot \Delta u(q_L) < 0.$$

3.3 Neutral Network and No Price Discrimination

Now consider a neutral network where the ISP is constrained to choose only a single price-quality pair (p, q). Given this single quality offer constraint, the ISP decides between serving only the high type CPs with the exclusion of the 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 $p = \alpha H u(q_H^{FB})$, which gives $\widetilde{\Pi}^{EX} = \underline{u} + \nu[H u(q_H^{FB}) - c q_H^{FB}]$.[17]

If the monopolistic ISP decides to serve both types, then $p = \alpha L u(q)$ and $f = \underline{u} + (1 - \alpha) \sum_\theta \nu_\theta \theta u(q)$. Hence, the monopoly ISP chooses a single quality q to solve

$$\max_q \ \Pi(\alpha, q) = \underline{u} + (L + (1 - \alpha)\nu\Delta)u(q) - cq.$$

From the first-order condition, we obtain the optimal quality choice when both types of CPs are served:

$$(L + (1 - \alpha)\nu\Delta)u'(\widetilde{q}(\alpha)) = c. \tag{6}$$

Equation (6) indicates that the optimal quality choice lies between the first-best level qualities for the high and the low types, that is, $q_L^{FB} \leq \widetilde{q}(\alpha) < q_H^{FB}$. If the ISP could extract CPs' surplus only, then $\widetilde{q}(\alpha)$ could never be higher than q_L^{FB}. However, we find that $\widetilde{q}(\alpha)$ is strictly higher than q_L^{FB} for any $\alpha < 1$. This has to do with the two-sided nature of the ISP's business: since it can extract extra consumer surplus generated by high type CPs in addition to $L u(q)$, it chooses a quality level higher than q_L^{FB}. By total differentiation of (6), we can derive that the quality decreases with α:

$$\frac{d\widetilde{q}(\alpha)}{d\alpha} = \frac{\nu\Delta u'(\widetilde{q})}{(L + (1 - \alpha)\nu\Delta)u''(\widetilde{q})} < 0. \tag{7}$$

Let $\widetilde{\Pi}(\alpha) \equiv \Pi(\alpha, \widetilde{q}(\alpha))$. From the envelope theorem, the monopolistic ISP's profit without exclusion strictly decreases with α as in the non-neutral regime.

$$\frac{d\widetilde{\Pi}(\alpha)}{d\alpha} = -\nu\Delta u(\widetilde{q}(\alpha)) < 0$$

By contrast, the ISP's profit under exclusion, $\widetilde{\Pi}^{EX}$, is independent of α. We assume

---

[17]We use a tilde (~) to denote variables associated with a neutral network.

the following.

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

This assumption is made to highlight the difference between one-sided markets and two-sided markets in evaluating the relative merit of no price discrimination (PD) vis-à-vis second-degree PD: even if second-degree PD generates higher welfare than no PD in one-sided markets, it does not necessarily mean that the same holds for two-sided markets. Assumption 2 is introduced to make this point clear.[18]   In addition, Assumption 2 is introduced to reduce the number of cases to consider.   If relevant, we will comment on how the details of the analysis can change when this assumption is violated.

Assumption 2, together with the monotonicity of $\widetilde{\Pi}(\alpha)$, implies that there exists a unique threshold level of α denoted by $\alpha^N \in (0,1)$ such that the monopolistic ISP serves both types of CPs for $\alpha < \alpha^N$ and excludes the low type CPs for $\alpha > \alpha^N$, where α^N is implicitly defined by $\widetilde{\Pi}(\alpha^N) = \widetilde{\Pi}^{EX}$, that is,

$$(L + (1 - \alpha^N)\nu\Delta)u(\widetilde{q}(\alpha^N)) - c\widetilde{q}(\alpha^N) = \nu \cdot (Hu(q_H^{FB}) - cq_H^{FB}). \tag{8}$$

Therefore, the monopolist ISP's profit under the neutral system, $\widetilde{\Pi}^M(\alpha)$, can be written as

$$\widetilde{\Pi}^M(\alpha) = \begin{cases} \widetilde{\Pi}(\alpha) \text{ for } \alpha < \alpha^N \\ \widetilde{\Pi}^{EX} \ \text{ for } \alpha \geq \alpha^N \end{cases}$$

and can be shown as in Figure 1.

Let $(\widetilde{p}^*(\alpha), \widetilde{q}^*(\alpha))$ represent the ISP's choice under a neutral network. Then, the quality chosen by the ISP is given by:

$$\widetilde{q}^*(\alpha) = \begin{cases} \widetilde{q}(\alpha) \text{ for } \alpha < \alpha^N \\ q_H^{FB} \text{ for } \alpha \geq \alpha^N \end{cases}$$

The corresponding retail prices are given by $\widetilde{p}^*(\alpha) = \alpha Lu(\widetilde{q}(\alpha))$ for $\alpha < \alpha^N$ and $\widetilde{p}^*(\alpha) = \alpha Hu(q_H^{FB})$ for $\alpha > \alpha^N$.[19]

---

[18]Note that Assumption 2 is not necessary if our purpose is only to show the possibility that net neutrality may be superior to non-neutrality. If Assumption 2 is violated, $\widetilde{\Pi}(\alpha = 1) > \widetilde{\Pi}^{EX}$ (i.e., no exclusion occurs under a neutral network for $\alpha = 1$). In such a case, a sufficient condition for the result is that a neutral network generates a strictly higher welfare than a non-neutral network for $\alpha = 1$. Then, by the continuity argument, the neutral network generates a higher welfare than the non-neutral network for α close to 1.

[19]If Assumption 2 is violated and $\widetilde{\Pi}(\alpha = 1) > \widetilde{\Pi}^{EX}$, there will be no exclusion of CPs. We thus have $\widetilde{\Pi}^M(\alpha) = \widetilde{\Pi}(\alpha)$, $\widetilde{q}^*(\alpha) = \widetilde{q}(\alpha)$, and $\widetilde{p}^*(\alpha) = \alpha Lu(\widetilde{q}(\alpha))$ for all α.
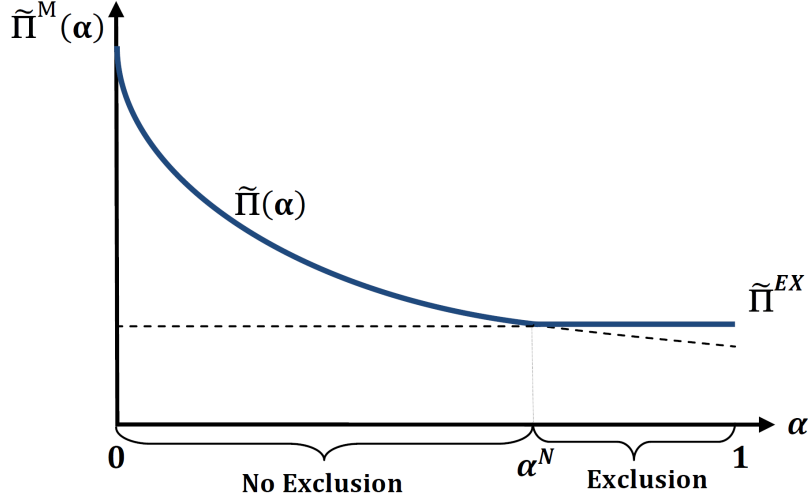
Figure 1: The monopolist ISP's profit in the neutral network

3.4 Assumptions and Social Welfare with One-sided Market

Under Assumptions 1-2, the non-neutral network dominates the neutral network, from the social welfare point of view, for the extreme cases of $\alpha = 1$ and $\alpha = 0$. Essentially, these two cases can be considered as representations 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 $\underline{u}$ only. So, the monopoly ISP will set $f = \underline{u}$ 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 in one-sided markets. In this case, high type CPs consume q_H^{FB} in both regimes, but low types are served only under a non-neutral network. This is a standard argument in favor of 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 a 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. 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:

$$u^{FB} = \sum_{\theta} \nu_\theta \theta u(q_\theta^{FB}), \tag{9}$$

$$c^{FB} = \sum_{\theta} \nu_\theta c q_\theta^{FB}. \tag{10}$$

14

Under a non-neutral regime, the monopoly ISP provides the first-best quality for each type of CPs and charges the consumer subscription fee $f(\alpha = 0) = \underline{u} + u^{FB}$.[20] By contrast, the monopoly ISP is constrained to offer one level of quality under a neutral regime and hence can never achieve the first-best outcome. In summary, we have:

Proposition 1. *Consider a monopoly ISP facing homogeneous consumers with inelastic subscription.*

(i) *If $\alpha = 1$, under Assumptions 1-2, the ISP serves both types of CPs in a non-neutral network but serves only high types in a neutral network. Therefore, social welfare is higher under a non-neutral network than under a neutral network.*

(ii) *If $\alpha = 0$, the outcome chosen by the ISP coincides with the first-best under a non-neutral network. By contrast, under a neutral network, the first-best can never be realized. 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. This result will be contrasted to the case where a neutral network can provide a higher social welfare relative to a non-neutral network, as we consider intermediate values of α.

The parameter α represents the surplus division between CPs and end users when they interact through the ISP and indicates which side the ISP should focus on to extract rents. As α increases, CPs capture more surplus and the extraction of rents from the CP side becomes more important. As a result, the ISP distorts the quality for low type CPs further down to reduce the rent of the high type CPs under the non-neutral network. Under the neutral network, ISPs exclude the low type CPs when α is high enough (i.e., $\alpha > \alpha^N$) to reduce the rents of high type CPs under Assumption 2. This finding may have important policy implications. For instance, α may capture how much CPs can extract consumer surplus through micropayments. From this perspective, the concern about potential exclusion of CPs in a neutral network can be heightened if the business model of CPs shifts from an advertising-based one with free access to the one with micropayments that directly charges consumers for content.

---

[20]When $\alpha = 0$, every CP makes zero profit and we can assume that a CP follows the ISP's desire in the case of indifference. 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 $p = \alpha H u(q)$.

So far, we have treated α as a parameter. In Appendix B, we provide a very simple informative advertising model as a micro-foundation to illustrate how α can be endogenized and affected by the development of targeted advertising. For instance, it is shown that if targeted advertising induces more effective advertising but solicits more privacy concerns due to its intrusiveness, α increases.

4 Comparison of Neutral vs. Non-neutral Networks

In this section, we compare quality choices and social welfare in the neutral network to those in the non-neutral network.

4.1 Quality Choices and Each Group's Payoff

Figure 2 shows the optimal quality schedules for both network regimes. In a non-neutral network, there is no distortion in the quality for high type CPs and a downward distortion in the quality for low type CPs. As α decreases, this distortion becomes smaller and becomes zero when $\alpha = 0$ (i.e., $q^*(\alpha = 0) = q_L^{FB}$). In a neutral network, the ISPs serve only high type CPs for $\alpha > \alpha^N$ and choose $\widetilde{q}^*(\alpha) = q_H^{FB}$; for $\alpha \leq \alpha^N$, they choose a quality $\widetilde{q}^*(\alpha) \in \left(q_L^{FB}, q_H^{FB}\right)$ and serve both types where $\widetilde{q}^*(\alpha)$ decreases with α.

The ISP realizes a strictly higher profit in a non-neutral network than in a neutral network by the revealed preference argument; in a non-neutral network the ISPs could always choose an equal quality for both delivery services if this would give a higher profit. Consumers are indifferent across the two regimes because their surplus is completely extracted by the ISP anyway.[21]

Low type CPs always receive zero rents for any α regardless of net neutrality regulation. Comparison of high type CPs' payoff depends on whether α is higher or lower than α^N. If $\alpha \leq \alpha^N$, the relationship of $\widetilde{q}^*(\alpha) > q_L^{FB} \geq q_L^*(\alpha)$ implies that high type CPs obtain a higher payoff in the neutral network than in the non-neutral network. If $\alpha > \alpha^N$, the reverse holds since high type CPs obtain no rent in the neutral network while they obtain a strictly positive rent in the non-neutral network.

---

[21]However, this result is due to the assumption of homogeneous consumers. In an extension with competing ISPs and heterogeneous consumers, we show that consumer surplus is higher in the non-neutral network than in the neutral network.
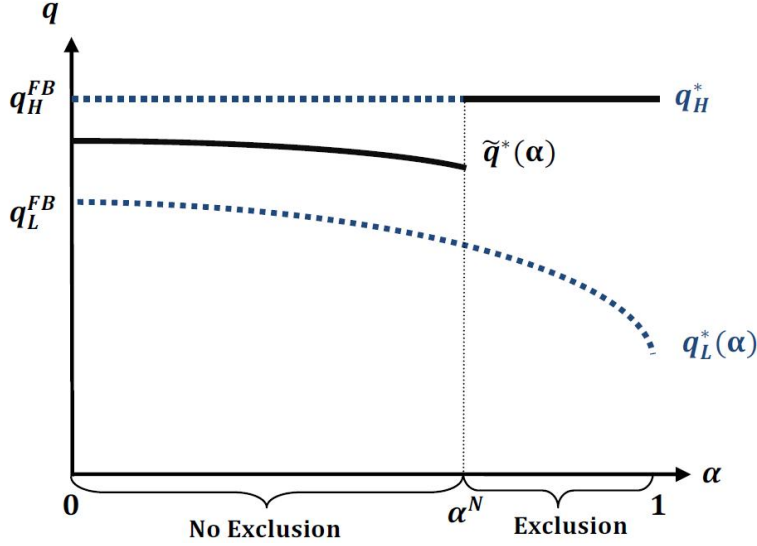
Figure 2: The optimal quality schedules

4.2 Social Welfare

We now perform a welfare comparison to assess the merit of net neutrality regulations that prohibit price discrimination in two-sided markets. Recall that under Assumptions 1 and 2, a neutral network cannot outperform a non-neutral network for the two extreme cases of full or zero extraction of surplus by CPs vis-à-vis consumers (see Proposition 1). We investigate whether this result is robust to intermediate cases of $\alpha \in (0,1)$ and find that the social welfare ranking between the two regimes can be reversed for intermediate values of α.

As we provided the intuition for this result in the introduction, the single quality restriction can be welfare-enhancing because of its smaller quality distortion in a neutral network compared to a non-neutral network, despite offering a suboptimal quality for the high type CPs. Note that the ISP has two sources of revenues: the one from the CP side and the one from the consumer side. When the ISP chooses the quality for low type CPs, it faces a trade-off which arises from the two-sided nature of the ISP's business. More precisely, a downward distortion in the quality for low type CPs has the benefit of extracting more rent from high type CPs, on the content side, and the cost of extracting less consumer surplus, on the consumer side. This implies that when the ISP focuses on making revenue from the consumer side rather than from the CP side, there will be less distortion in quality. *Because the ISP has more instruments to extract CPs' surplus in a non-neutral network than in a neutral network, it focuses relatively more on extracting CPs' surplus in a non-*

neutral network than in a neutral network. In our model, this arises especially when α becomes smaller than α^N. For $\alpha \in (0, \alpha^N)$, the ISP always downwardly distorts the quality for low type CPs in a non-neutral network whereas in a neutral network, it serves both types of CPs with a quality superior to q_L^{FB}. For this reason, a neutral network may provide a higher social welfare than a non-neutral network.[22]

Given this insight, we now provide more rigorous mathematical derivation for our result. The social welfare in the non-neutral network with optimal quality choices can be written as

$$W^* = \underline{u} + \sum_\theta \nu_\theta [\theta u(q_\theta^*) - c q_\theta^*] = \underline{u} + \nu (H u(q_H^{FB}) - c q_H^{FB}) + (1-\nu)(L u(q_L^*(\alpha)) - c q_L^*(\alpha)). \tag{11}$$

We take the first-order derivative of the social welfare with respect to α as

$$\frac{dW^*}{d\alpha} = \frac{d[(1-\nu)(L u(q_L^*(\alpha)) - c q_L^*(\alpha))]}{d\alpha}. \tag{12}$$

Using the first-order optimal quality condition for the low type CPs in (4), we can rewrite (12) as follows:

$$\frac{dW^*}{d\alpha} = \alpha \nu \Delta u'(q_L^*) \frac{dq_L^*}{d\alpha} < 0. \tag{13}$$

The inequality above holds because the quality distortion increases in α, i.e., $\frac{dq_L^*}{d\alpha} < 0$.

Similarly, we can define social welfare under the neutral network and, for the same reason as in the non-neutral network, we find that the social welfare in the neutral network also decreases in α for any $\alpha < \alpha^N$:

$$\frac{d\widetilde{W}^*}{d\alpha} = \alpha \nu \Delta u'(\widetilde{q}^*) \frac{d\widetilde{q}^*}{d\alpha} < 0. \tag{14}$$

Recall that the quality adjustment to the change in α can be derived as (5) for the non-neutral network and (7) for the neutral one.

To facilitate the comparison further and gain more intuition, let us consider a 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 with positive

---

[22]If we relax Assumption 2 and consider the case where a neutral network entails no exclusion and the quality distortion effect is high enough in a non-neutral network, social welfare may be higher in a neutral network when α is equal to 1 (and close to 1, by continuity).

constants A and B. Then, we can obtain a clear comparison of

$$\left| \frac{dq_L^*(\alpha)}{d\alpha} \right| > \left| \frac{d\widetilde{q}^*}{d\alpha} \right| \quad \text{for } \forall \nu \in (0,1) \tag{15}$$

from $-\frac{u'(\widetilde{q}^*)}{u''(\widetilde{q}^*)} = -\frac{u'(q_L^*)}{u''(q_L^*)} = r$. This implies that the ISP's quality degradation gradient for low types in the non-neutral network is steeper than the one for the uniform quality in the neutral network as α increases. In addition, we find $u'(q_L^*) > u'(\widetilde{q}^*)$ from $q_L^* < \widetilde{q}^*$ for any utility function with $u'' < 0$. Hence, we find that the social welfare decreases more quickly as α increases in the non-neutral network compared to the neutral network, i.e., $\left| \frac{dW^*}{d\alpha} \right| > \left| \frac{d\widetilde{W}^*}{d\alpha} \right|$.

Given this understanding, let us finally compare the level of social welfare under two different network regimes. Recalling the definition of α_N in (8), social welfare in the neutral network at $\alpha = \alpha^N$ can be expressed as

$$\begin{aligned}
\widetilde{W}^* \Big|_{\alpha=\alpha^N} &= \underline{u} + (L + (1 - \alpha^N)\nu\Delta)u(\widetilde{q}^*) - c\widetilde{q}^* + \alpha^N \nu \Delta u(\widetilde{q}^*) \\
&= \underline{u} + \nu \cdot (Hu(q_H^{FB}) - cq_H^{FB}) + \alpha^N \nu \Delta u(\widetilde{q}^*)
\end{aligned}$$

This simplifies the comparison between $\widetilde{W}^*$ and W^* evaluated at $\alpha = \alpha^N$ as the comparison between $\alpha^N \nu \Delta u(\widetilde{q}^*)$ and $(1-\nu)(Lu(q_L^*) - cq_L^*)$:

$$\widetilde{W}^* \Big|_{\alpha=\alpha^N} > W^*|_{\alpha=\alpha^N} \Leftrightarrow \alpha^N \nu \Delta u(\widetilde{q}^*) > (1-\nu)(Lu(q_L^*) - cq_L^*) \tag{16}$$

Under Assumptions 1-2, social welfare can be higher in a neutral network relative to a non-neutral network as long as (16) is satisfied. Since $\widetilde{W}^* \Big|_{\alpha=0} < W^*|_{\alpha=0}$ and $\left| \frac{dW^*}{d\alpha} \right| > \left| \frac{d\widetilde{W}^*}{d\alpha} \right|$, if (16) is satisfied, there exists a unique α^* such that $\widetilde{W}^* > W^*$ for any $\alpha \in (\alpha^*, \alpha^N]$.

Proposition 2. *Consider the CARA utility function for which Assumptions 1 and 2, and (16) are satisfied. There exists a unique level of α^*, which is weakly smaller than α^N but is greater than zero, such that social welfare is higher in the neutral network than in the non-neutral network for all $\alpha \in (\alpha^*, \alpha^N]$.*

Figure 3 illustrates a plausible case in which a neutral network may yield a higher total social welfare than a non-neutral network. As a numerical exercise, consider parameters such that $H = 40$, $L = 30$, $\underline{u} = 5$, and $\nu = 0.74$. If we consider a CARA utility function such as $u(q) = 1 - \frac{1}{2}e^{-2q}$, it satisfies all assumptions that we make and the neutral network yields the higher social welfare for $\alpha \in (0.2441, 0.8711)$ than
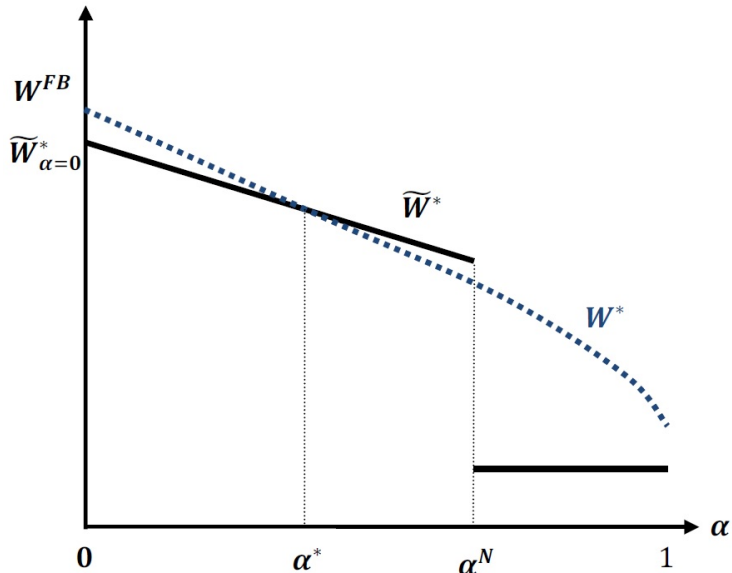
Figure 3: Welfare under neutral and non-neutral networks

the non-neutral network.

A simple application of the implicit function theorem to equation (8) that defines the critical value α^N yields the following comparative statics results.

Corollary 1. *(a)* $\frac{\partial \alpha^N}{\partial c} < 0$; *(b)* $\frac{\partial \alpha^N}{\partial \nu} < 0$

As Corollary 1 shows, the exclusion strategy is more likely to occur when the marginal cost of delivery increases and the proportion of high-type CP increases, with all other things being equal. This result has some implications for mobile Internet networks that are constrained by the scarcity of bandwidth imposed by physical laws and thus have a higher delivery cost (i.e., higher c) compared to fixed Internet networks with fiber optic cables; the non-neutral network is likely to increase the allocative efficiency and may provide justifications for differential treatments of mobile networks.

Though we frame our discussion in the context of the recent debate on network neutrality in the Internet, our analysis can be more generally interpreted. In particular, our findings suggest that welfare implications of second-degree price discrimination in two-sided platform markets can crucially depend on the relative allocation of the total surplus between the two sides.

Table 1 lists the top ten companies in terms of bandwidth consumption, their corresponding precentage of peak US Internet traffic (as of Feb 3, 2014) and business models.[23] Not surprisingly, popular streaming content providers in music, video,

---

[23]The original source for the traffic shares is DeepField. See Fitzgerald and Wakabayashi (2014).

Company	Service Type	Primary Business Models
Netflix (32.0 %)	video streaming	subscription fees
Google (22.0%)	search, video-streaming (Youtube)	ads
Apple (4.3%)	iTunes, App Store, iCloud	content sales, ads
Twitch (1.8%)	live streaming service for games	ads, subscription fees
Hulu (1.7%)	video streaming	subscription fees
Facebook (1.5%)	social networking	ads
Valve(1.3%)	online game platform	pricing per game
Amazon (1.2%)	e-commerce, video streaming (Prime)	retailing, subscription fees
Pandora (0.5%)	music streaming	ads
Tumblr (0.4%)	microblogging, social networking	ads

Table 1: Top Ten US Internet Sites in Bandwidth Consumption

and games make the list. As major CPs rely more on subscription-based services and develop better targeted advertising, we expect the parameter α in our model to increase over time. To the extent that these developments do not result in the exclusion of low type CPs under the neutral network, there should be more concern for the ISP's menu pricing on the CP side as they induce a greater downward quality distortion for low type CPs.

5   Competing ISPs with Interconnection

We have so far analyzed the effects of price discrimination in two-sided markets with a monopolistic ISP. In this section, we show the robustness of our results in the framework of competing ISPs with interconnection and heterogeneous consumers. This section can be of independent interest as a contribution to the literature on interconnection. In particular, we establish the off-net cost pricing principle in the presence of second-degree price discrimination and show how a reciprocal access charge agreement can be used to replicate the monopolistic outcome between competing ISPs in a two-sided market.

5.1   The Model

In order to model competition between ISPs with elastic demand, we consider a discrete choice model on the consumer side. There are two interconnected ISPs denoted by $i = 1, 2$. Consumers choose one of the three options: subscribing to either one of the two ISPs or subscribing to neither of them. A consumer k's gross utility from subscribing to ISP i is represented by $U_i(\alpha) - \epsilon_{ik}$, where $\epsilon_{ik} \in [0, \infty)$. The

component $U_i(\alpha)$ is common to all consumers. The idiosyncratic consumer-specific component ϵ_{ik} reflects heterogeneity in consumer preferences towards horizontally differentiated ISPs. We assume that ϵ_{1k} and ϵ_{2k} are independently and identically distributed with a strictly positive density. As in the monopoly case, we have $U_i(\alpha)$ $= \underline{u} + (1 - \alpha)\sum_\theta \nu_{\theta i}\theta u(q_\theta)$, where $\nu_{\theta i}$ is the measure of type θ CPs whose content can be consumed by subscribing to ISP i. Note that q_θ is not specific to ISP i since we consider cooperative choice of quality. Under interconnection with reciprocal access pricing, we have $\nu_{\theta 1} = \nu_{\theta 2} \equiv \nu_\theta^*$ where ν_θ^* is the measure of type θ CPs subscribed to any of the two ISPs. The common component in a consumer's net utility from subscribing to ISP i, u_i, is given by

$$u_i(\alpha, f_i) = U_i(\alpha) - f_i, \tag{17}$$

where f_i is the subscription price charged by ISP i.

With the total number of potential consumers normalized to 1, the subscription demand for ISP i can be described as

$$n_i(u_i, u_j) = \Pr(\epsilon_i - \epsilon_j < u_i - u_j \text{ and } \epsilon_i < u_i). \tag{18}$$

We assume that $n_i(.)$ is twice differentiable. Our demand specification implies that $\frac{\partial n_i}{\partial u_i} > 0$, $\frac{\partial n_j}{\partial u_i} < 0$ for $i \neq j$ and $\frac{\partial(n_i + n_j)}{\partial u_i} > 0$.[24]

To model interconnection between the two ISPs, we assume that the total marginal cost has two components, i.e., $c = c_O + c_T$, where $c_O \geq 0$ and $c_T \geq 0$ stand for the cost of origination and that of termination per quality, respectively. For simplicity, we consider cooperative choice of quality and access charge under both regimes. Under a non-neutral regime, ISPs 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 for low type CPs. The termination charge *per unit quality* for type θ traffic can be implicitly defined as a_θ. 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 $\widehat{c}_\theta \equiv c + a_\theta - c_T (= c_O + a_\theta)$ denote the perceived unit quality cost of the

---

[24]In an earlier version of this paper, we adopted a "Hotelling model with hinterlands" (Armstrong and Wright, 2009 and Hagiu and Lee, 2011) to illustrate our results. The Hotelling model with hinterlands can be considered as a special case of our general framework.

off-net content that terminates in the other network for quality q_θ.[25] Let $\widehat{\mathbf{c}} \equiv (\widehat{c}_H, \widehat{c}_L)$ in the non-neutral network.

In a neutral regime, ISP i is constrained to offer a single uniform delivery quality q.[26] The ISPs jointly choose a single quality level and a single per unit access charge a. Let $\widehat{c} \equiv c + a - c_T$ denote the off-net cost per unit quality in the neutral network.

We note that because of the interconnection agreement between the two ISPs, a CP can reach any consumer subscribed to either ISP regardless of the ISP it chooses to deliver its content.

The game is played in the following sequence.

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

- Stage 2: In the non-neutral regime, each ISP i with $i = 1, 2$ simultaneously sets for CPs $\{p_i(q_H), p_i(q_L)\}$, a menu of prices per unit delivery of content of given quality. In the neutral regime, there is only one delivery quality and each ISP sets a price of $p_i(q)$. Given the price schedules, each CP decides whether to participate in the market, and if it participates, decides which ISP to use to deliver its content (and what type of delivery service to purchase in the non-neutral regime).

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

One main reason to consider this sequential timing rather than two alternative timing scenarios where stages 2 and 3 are reversed or take place simultaneously is that the ISPs have less incentive to deviate from the joint-profit maximizing prices under this sequential timing than under the other ones.[27] We establish that the off-net cost pricing on the content side should hold in equilibrium regardless of the timing and

---

[25]In a non-neutral regime, we can further distinguish two cases depending on whether or not termination-based price discrimination (TPD) is possible. With TPD, ISP i proposes a pricing 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). In this paper, we do not consider the possibility of TPD, that is, we analyze only the case where $p_i(q) = \widehat{p}_i(q)$. However, the qualitative results do not change when we consider TPD if CPs are allowed to multi-home.

[26]Under the neutral regime, there is no TPD because content cannot be treated differently depending on its destination.

[27]This is because off-net cost pricing is a necessary condition but not a sufficient condition for an equilibrium. For instance, under simultaneous timing, if ISP i deviates in its offer to CPs, it

show that there is an upper bound on the ISPs' joint profit associated with the off-net cost pricing. Hence, the upper bound does not depend on the timing we choose. Finally, we show that the ISPs can achieve this upper bound under the sequential timing specified above.

5.2 Competition in CP Market and Off-Net Cost Pricing

LMRT (2003) first showed that in a broad range of environments, network operators set prices for their customers as if their customers' traffic were entirely off-net, which they termed the off-net cost pricing principle. We extend their analysis and confirm that their result is robust to the introduction of heterogeneous content types with menu pricing and to alternative timing assumptions.

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

Proof. See the Appendix. □

Lemma 1 shows that off-net cost pricing is a necessary condition that any equilibrium price for CPs generating positive sales must satisfy. This property holds more generally regardless of the timing we consider. In fact, it is straightforward to prove it for the sequential timing of reverse order (i.e., stages 2 and 3 are reversed) or for the simultaneous timing (i.e., stages 2 and 3 take place at the same time). With these alternative timing assumptions, a necessary condition that equilibrium prices on the content side should satisfy is that an ISP should be indifferent between winning and losing a given type of CPs, given the prices and subscription decisions on the consumer side.

For instance, consider a neutral network and let $p(q)$ be an equilibrium price for CPs given that the ISPs previously agreed on (q, a). We normalize the total number of consumers subscribed to one, without loss of generality, and let $s_i \in [0, 1]$ represent ISP i's consumer market share. Suppose first that at $p(q)$ both types of CPs buy connections from ISP i. Then, ISP i's profit from the content side in equilibrium is $p(q) - s_i c q - (1 - s_i)(c + a - c_T)q$. If it loses the CPs by charging a higher price, its profit from the content side will be $s_i(a - c_T)q$. Therefore, the following inequality

---

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 its offer to CPs, ISP j can adjust its offer to consumers in stage 3.

must hold in equilibrium:

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

which is equivalent to

$$p(q) \geq (c + a - c_T)q = \widehat{c}q.$$

Symmetrically, the condition for ISP j to weakly prefer not serving any CPs to attracting both types of CPs gives the condition

$$p(q) \leq (c + a - c_T)q = \widehat{c}q.$$

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

Suppose now that at $p(q)$, only high type CPs buy connection and ISP i wins them. Then, ISP i's equilibrium profit from content is $\nu[p(q) - s_i cq - (1 - s_i)(c + a - c_T)q]$ and its content side profit from losing the 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 a non-neutral network.

We also would like to point out that even though off-net cost pricing is a necessary condition for an equilibrium, it is not a sufficient condition for any arbitrary access charge because an ISP may have an incentive to deviate.[28] However, we later on show that it is both necessary and sufficient for the access charge(s) optimally agreed on by the ISPs (to maximize their joint profits) in stage 1.

Now we examine the profit that each ISP obtains on the content side under the off-net cost pricing. Let n_i be the expected number of consumers subscribed to ISP i for $i = 1, 2$ at stage 3. Consider a given CP who uses quality q under the off-net cost

---

[28]To illustrate this point, consider a neutral network and the sequential timing of reverse order (i.e., stage 2 and stage 3 are reversed). Suppose that $a = (\alpha Lu(q) + \varepsilon)/q - c_O$ where $\varepsilon > 0$ is infinitesimal. Then, off-net cost pricing leads to $p(q) = \alpha Lu(q) + \varepsilon$. Hence, only high types purchase the quality at off-net cost pricing. Consider now the deviation of ISP i to $p'(q) = \alpha Lu(q)$ such that both types purchase the quality. This deviation is profitable in the content side if and only if

$$\nu s_i(\alpha Lu(q) + \varepsilon - cq) < \alpha Lu(q) - s_i cq - (1 - s_i)(\alpha Lu(q) + \varepsilon),$$

where s_i is ISP i's market share in the content side at the off-net cost. The condition is equivalent to

$$[\nu s_i + (1 - s_i)]\varepsilon < (1 - \nu)s_i(\alpha Lu(q) - cq),$$

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

pricing $\widehat{c}q$. If this CP subscribed to ISP i at stage 2, then ISP i's profit from this CP is

$$\widehat{c}q(n_i + n_j) - cqn_i - (c + a - c_T)\,qn_j = (\widehat{c} - c)qn_i.$$

If this CP subscribed to ISP j at stage 2, then ISP i's profit from this CP is

$$(a - c_T)qn_i = (\widehat{c} - c)qn_i.$$

Therefore, we derive:

Lemma 2 (Profit from CPs). *Consider any off-net cost pricing equilibrium. Then, ISP i's profit from the CP side is given by $n_i\widehat{\pi}^{CP}$, where n_i is the expected number of consumers subscribed to ISP i for $i = 1, 2$ and $\widehat{\pi}^{CP} \equiv \sum_\theta \nu_\theta^*(a_\theta - c_T)q_\theta = \sum_\theta \nu_\theta^*(\widehat{c}_\theta - c)q_\theta$ where ν_θ^* is the measure of type θ CPs that subscribed to any of the two ISPs. This result holds regardless of whether networks are neutral or not. In the neutral network, if both types are served, it is required that $a_H = a_L = a$.*

Note that the result of this lemma does not depend on the subscription distribution of CPs across the two ISPs because each ISP is indifferent between winning and losing CPs at off-net cost pricing. Note that $\widehat{\pi}^{CP}$ represents the profit *per consumer* that each ISP makes from the content side with off-net cost pricing and does not depend on (n_1, n_2).

5.3 Competition in Consumer Market: the Equivalence Result

Given $(q_\theta; a_\theta)$ which the ISPs agreed upon at stage 1, assume that the off-net cost pricing holds at stage 2. Then, in any symmetric equilibrium, each ISP's profit can be written as $\Pi = n \cdot (f + \widehat{\pi}^{CP})$ from Lemma 2. Because the number of subscribers depends on consumer net surplus $U(\alpha) - f$, the equilibrium subscription fee f is a function of per consumer utility $U(\alpha)$ and CP profit per consumer $\widehat{\pi}^{CP}$. Thus, each ISP's profit in a symmetric equilibrium can be written as

$$\Pi(U(\alpha), \widehat{\pi}^{CP}) = n(U(\alpha) - f(U(\alpha), \widehat{\pi}^{CP})) \cdot [f(U(\alpha), \widehat{\pi}^{CP}) + \widehat{\pi}^{CP}]. \qquad (19)$$

We show that $\Pi(U(\alpha), \widehat{\pi}^{CP})$ attains the maximum value when $(U(\alpha) + \widehat{\pi}^{CP})$ is maximized. Notice that $(U(\alpha) + \widehat{\pi}^{CP})$ evaluated at $p_\theta = \widehat{c}_\theta$ is exactly the total profit of the monopolistic ISP studied in Sections 3-4, that is, $\Pi^M(\alpha) = U(\alpha) + \widehat{\pi}^{CP}$. We thus establish the following *equivalence result* that the maximization of the joint profit of the competing ISPs requires maximizing $\Pi^M(\alpha)$ with respect to $(q_\theta; a_\theta)$ under non-

neutral networks (respectively, $(q; a)$ under neutral networks). Mathematically, it requires the ISPs' joint profit increases in Π^M, i.e., $\frac{d\Pi}{d\Pi^M} > 0$.

We find that the equivalence result crucially depends on the *pass-through* rates of $U(\alpha)$ and $\widehat{\pi}^{CP}$ into f: how much of an increase in consumer valuation from content delivery $U(\alpha)$ or in CP profit per consumer $\widehat{\pi}^{CP}$ is translated into a change in the equilibrium subscription fee, i.e., $\partial f / \partial U(\alpha)$ and $\partial f / \partial \widehat{\pi}^{CP}$. We expect that $\partial f / \partial U(\alpha) > 0$, but $\partial f / \partial \widehat{\pi}^{CP} < 0$ because a higher profit from the content side induces fiercer competition on the single-homing consumer side which constitutes the competitive bottleneck (Armstrong, 2006). We consider a typical scenario in two-sided markets where the pass-through rates are not full, i.e., $0 < \partial f / \partial U(\alpha) < 1$ and $-1 < \partial f / \partial \widehat{\pi}^{CP} < 0$. The partial pass-through condition is guaranteed by the following assumption.

Assumption 3. $\frac{\partial^2 n_i}{\partial u_i^2} + \frac{\partial^2 n_i}{\partial u_i \partial u_j} \leq 0$.

Assumption 3 states that the Jacobian of the demand system has a dominant diagonal. This is a standard assumption to guarantee the uniqueness of equilibrium in Bertrand competition among differentiated products (Milgrom and Roberts, 1990; Vives, 2000). The following relationship proves useful in establishing our equivalence result.

Lemma 3 (Pass-through rates). *(i)* $\frac{\partial f}{\partial U(\alpha)} + \left| \frac{\partial f}{\partial \widehat{\pi}^{CP}} \right| = 1$

(ii) Under Assumption 3, $\left(\frac{\partial f}{\partial U(\alpha)}, -\frac{\partial f}{\partial \widehat{\pi}^{CP}} \right) \in (0, 1)^2$.

Proof. See the Appendix. □

Since $\Pi = n \cdot (f + \widehat{\pi}^{CP})$, a sufficient condition for $\frac{d\Pi}{d\Pi^M} > 0$ is to have both n and $(f + \widehat{\pi}^{CP})$ increase in $\Pi^M(\alpha) = U(\alpha) + \widehat{\pi}^{CP}$:

$$\frac{d\left[U(\alpha) - f \right]}{d\Pi^M} > 0 \text{ and } \frac{d[f + \widehat{\pi}^{CP}]}{d\Pi^M} > 0, \tag{20}$$

which is equivalent to

$$-\frac{d\widehat{\pi}^{CP}}{d\Pi^M} < \frac{\partial f}{\partial U(\alpha)} \frac{dU(\alpha)}{d\Pi^M} + \frac{\partial f}{\partial \widehat{\pi}^{CP}} \frac{d\widehat{\pi}^{CP}}{d\Pi^M} < \frac{dU(\alpha)}{d\Pi^M}. \tag{21}$$

From $\frac{dU(\alpha)}{d\Pi^M} + \frac{d\widehat{\pi}^{CP}}{d\Pi^M} = 1$, the condition above can be rewritten as

$$\frac{dU(\alpha)}{d\Pi^M} - 1 < \frac{\partial f}{\partial U(\alpha)} \frac{dU(\alpha)}{d\Pi^M} + \left| \frac{\partial f}{\partial \widehat{\pi}^{CP}} \right| \left(\frac{dU(\alpha)}{d\Pi^M} - 1 \right) < \frac{dU(\alpha)}{d\Pi^M}. \tag{22}$$

27

By Lemma 3, the middle expression in the above inequalities is a convex combination of the LHS and RHS of the inequalities. Therefore, the sufficient condition for the equivalence result holds. Essentially, (20) tells us that each component of the equilibrium profit is an increasing function of Π^M if the direct positive effect from a change in $U(\alpha)$ or $\widehat{\pi}^{CP}$ is not overturned by the corresponding change in f. This occurs, according to (22), when the two pass-through rates are not full and the sum of their absolute values is equal to one.

Proposition 3 (The Equivalence Result). *Under the off-net cost pricing, the competing ISPs maximize the same objective as a monopoly ISP facing homogeneous consumers.*

5.4 ISPs' Choice of Quality and Access Charges

We showed that the competing ISPs maximize the same objective as a monopoly ISP facing homogeneous consumers and that off-net cost pricing must hold in any equilibrium. One potential issue is 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. However, we show that for the access charge optimally agreed on by the ISPs, this issue does not arise and off-net cost pricing is both necessary and sufficient for an equilibrium in the CP market.

We proceed in two steps. First, we consider a constrained benchmark case 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 in section 3 on the content side. This 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, $p(q_\theta) = \widehat{c}_\theta q = (c + a_\theta - c_T)q_\theta$. Second, we consider the original case in which any ISP is allowed to deviate from off-net cost pricing in stage 2 and prove that there is no profitable deviation when the ISPs agree on the qualities and access charges that would implement the monopoly benchmark outcome.

5.4.1 Non-Neutral Network

We first consider ISPs' choice of quality levels and access charges in the non-neutral network. According to the equivalence result, the ISPs collectively choose the quality levels and the corresponding access charges to maximize $U(\alpha) + \widehat{\pi}^{CP}$. From (3) and off-net cost pricing, it is immediate that the access charges will be chosen as follows

to replicate the monopolistic solution:

$$a_H^* = \alpha\left(Hu(q_H^*) - \Delta u(q_L^*)\right)/q_H^* - c_O \quad \text{and} \quad a_L^* = \alpha Lu(q_L^*)/q_L^* - c_O. \quad (23)$$

Proposition 4. *Consider a non-neutral network under Assumption 1:*

   (i) Suppose that no ISP is allowed to deviate from the off-net cost pricing.

        (a) 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 (4).

        (b) The ISPs choose access charges (a_H^*, a_L^*) given by (23).

   (ii) When the ISPs agree on (q_H^*, q_L^*) and (a_H^*, a_L^*), there is no profitable deviation from the off-net cost pricing and the ISPs implement the outcome that maximizes the joint profit.

    The proof of Proposition 4-(ii) is provided in the Appendix.

5.4.2   Neutral Network

As in the non-neutral network, the monopolistic ISP solution can be replicated by an appropriate choice 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 $\widetilde{q}^*(\alpha) = q_H^{FB}$ and the access charge of $\widetilde{a}^*(\alpha) = \alpha Hu(q_H^{FB})/q_H^{FB} - c_O$ to replicate the monopolistic solution. For $\alpha < \alpha^N$, the ISPs choose to serve both types of CPs with $\widetilde{q}^*(\alpha) = \widetilde{q}(\alpha)$ and the corresponding access charge of $\widetilde{a}^*(\alpha) = \alpha Lu(\widetilde{q}(\alpha))/\widetilde{q}(\alpha) - c_O$.

Proposition 5. *Consider a neutral network under Assumption 2.*

   (i) Suppose that no ISP is allowed to deviate from the off-net cost pricing. Then there exists a unique threshold level of α, denoted by $\alpha^N \in (0,1)$.

        (a) The ISPs offer $\widetilde{q}^*(\alpha) = q_H^{FB}$ for $\alpha > \alpha^N$ and $\widetilde{q}^*(\alpha) = \widetilde{q}(\alpha)$ otherwise.

        (b) The ISPs choose an access charge $\widetilde{a}^*(\alpha) = [\alpha Hu(q_H^{FB})/q_H^{FB}] - c_O$ for $\alpha > \alpha^N$ and $\widetilde{a}^*(\alpha) = [\alpha Lu(\widetilde{q}(\alpha))/\widetilde{q}(\alpha)] - c_O$ otherwise.

   (ii) When the ISPs agree on $\widetilde{q}^*(\alpha)$ and $\widetilde{a}^*(\alpha)$, there is no profitable deviation from the off-net cost pricing and the ISPs implement the outcome that maximizes the joint profit.

    The proof of Proposition 5-(ii) is provided in the Appendix.

5.5 Comparison of Social Welfare with Interconnection

Due to the equivalence result we have established, the quality choices in each regime and the effect of net neutrality regulation on each party's payoffs largely parallels those in the monopolistic ISP case. The social welfare result in the monopolistic ISP case also carries over to the case of competing ISPs with interconnection due to the equivalence result. The only difference is the effect on consumers. With heterogeneous consumers and competing ISPs, consumers' surplus cannot be completely extracted. In other words, we need to account for elastic consumer participation.

More specifically, given consumer k, define $\xi_k = \min[\epsilon_{1k}, \epsilon_{2k}]$ and let $G(.)$ and $g(.)$ respectively denote the corresponding cumulative and probability density functions of ξ_k.[29] Consider a symmetric equilibrium in a non-neutral network where consumers' common net utility level is given by u^*. Then, the number of consumers subscribing to ISPs is given by $N^* = G(u^*)$. The social welfare in the non-neutral network with optimal quality choices can be written as

$$\Omega^* = G(u^*) \cdot \omega^* - \int_0^{u^*} \xi dG \tag{24}$$

where ω^* the net social surplus per consumer (gross of the idiosyncratic component) at the optimal quality choices with $q_H^* = q_H^{FB}$ and $q_L^* = q_L^*(\alpha)$, that is,[30]

$$\omega^* = \underline{u} + \sum_\theta \nu_\theta [\theta u(q_\theta^*) - cq_\theta^*] = \underline{u} + \nu(Hu(q_H^{FB}) - cq_H^{FB}) + (1-\nu)(Lu(q_L^*(\alpha)) - cq_L^*(\alpha)).$$

We take the first-order derivative of the social welfare with respect to α as

$$\frac{d\Omega^*}{d\alpha} = G(u^*)\frac{\partial \omega^*}{\partial \alpha} + (\omega^* - u^*)G'(u^*)\frac{\partial u^*}{\partial \alpha}. \tag{25}$$

The first term in (25) is negative as we have shown in section 4. Its second term is also negative as $\omega^* - u^* > 0$ and $\frac{\partial u^*}{\partial \alpha} < 0$. Notice that $\omega^* - u^*$ is positive as long as $N^* = G(u^*)$ is smaller than the first-best level, which holds for any $\alpha \in [0, 1]$.[31] We have seen in the monopoly case that the ISP's profit decreases as α increases, which

---

[29] From the i.i.d. assumption on $(\epsilon_{1k}, \epsilon_{2k})$, neither $G(.)$ nor $g(.)$ depends on k.

[30] Note that the expression for ϖ^* is the same as W^* we defined in Eq. (11) when the size of the consumer market was fixed and normalized to 1.

[31] Even if $\alpha = 0$ and hence the ISPs choose the first-best qualities, N^* is smaller than the first-best level since the first-best outcome can be implemented only with zero profit of the ISPs.

implies that u^* decreases with α. More precisely,

$$\frac{\partial u^*}{\partial \alpha} = \frac{\partial [U - f(U, \widehat{\pi}^{CP})]}{\partial \alpha} = \frac{\partial U}{\partial \alpha} - \frac{\partial f}{\partial U}\frac{\partial U}{\partial \alpha} - \frac{\partial f}{\partial \widehat{\pi}^{CP}}\frac{\partial \widehat{\pi}^{CP}}{\partial \alpha}$$

Due to the relationship $\widehat{\pi}^{CP} = \Pi^M - U$ and Lemma 3, we can derive

$$\frac{\partial u^*}{\partial \alpha} = \frac{\partial U}{\partial \alpha}\underbrace{(1 - \frac{\partial f}{\partial U} + \frac{\partial f}{\partial \widehat{\pi}^{CP}})}_{(=0 \text{ by Lemma 3})} - \frac{\partial f}{\partial \widehat{\pi}^{CP}}\frac{\partial \Pi^M}{\partial \alpha} = (1 - \frac{\partial f}{\partial U})\frac{\partial \Pi^M}{\partial \alpha} < 0.$$

By using the first-order optimal quality condition for low type CPs and the expression for $\frac{\partial u^*}{\partial \alpha}$ above, we have

$$\frac{d\Omega^*}{d\alpha} = G(u^*) \cdot \alpha\nu\Delta u'(q_L^*)\frac{dq_L^*}{d\alpha} - (\omega^* - u^*)G'(u^*)(1 - \frac{\partial f}{\partial U})\frac{\partial \Pi^M}{\partial \alpha} < 0. \qquad (26)$$

Similarly, we can define social welfare under the neutral network and, for the same reason as in the non-neutral network, we find that the social welfare in the neutral network also decreases in α for any $\alpha < \alpha^N$:

$$\frac{d\widetilde{\Omega}^*}{d\alpha} = G(\widetilde{u}^*) \cdot \alpha\nu\Delta u'(\widetilde{q}^*)\frac{d\widetilde{q}^*}{d\alpha} - (\widetilde{\omega}^* - \widetilde{u}^*)\,G'(\widetilde{u}^*)(1 - \frac{\partial f}{\partial U})\frac{\partial \widetilde{\Pi}^M}{\partial \alpha} < 0. \qquad (27)$$

For the CARA utility function, we have shown that the ISPs' quality degradation gradient for low types in the non-neutral network is steeper than the one for the uniform quality in the neutral network as α increases and $u'(q_L^*) > u'(\widetilde{q}^*)$ for any utility function with $u'' < 0$. Since the non-neutral network provides consumer surplus at least as high as that under the neutral network, we have $G(u^*) > G(\widetilde{u}^*)$. Hence, we find that the social welfare decreases more quickly as α increases in the non-neutral network compared to the neutral network, i.e., $\left|\frac{d\Omega^*}{d\alpha}\right| > \left|\frac{d\widetilde{\Omega}^*}{d\alpha}\right|$ if the market expansion is highly limited ($G' \approx 0$) or the pass-through rate $\frac{\partial f}{\partial U}$ is close to one. We also know that per consumer social welfare can be higher in a neutral network relative to a non-neutral network as long as (16) is satisfied. Note that this does not ensure that a neutral network always dominates a non-neutral network at α^N because of $G(u^*) > G(\widetilde{u}^*)$. However, if G' is sufficiently small or the pass-through rate $\frac{\partial f}{\partial U}$ is close to one, the difference in number of consumers subscribed is of second-order relative to the difference in per consumer welfare. Hence, we can state that for a sufficiently small G' or a $\frac{\partial f}{\partial U}$ close to one, we have qualitatively the same welfare result as in the monopoly case even when we have competing ISPs with interconnection.

6 Concluding Remarks

In this paper we have analyzed the effect of net neutrality regulation when content is heterogeneous in its sensitivity to delivery quality. We consider 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 derive 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. We first establish the main results in a framework of a monopolistic ISP, which serves as a platform to connect CPs and end consumers. With interconnected networks, however, the assurance of a certain level of delivery quality requires cooperation among networks. To address this issue, we also have developed a model of interconnection in two-sided markets with competing ISPs and check the robustness of our results. Looking forward, this paper is a first step towards incorporating heterogeneous content in the analysis of interconnection issues.

There are many worthwhile extensions that call 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 constraints for ISPs and by considering a fixed mass of active CPs. The effects of net neutrality regulation on ISPs' capacity expansions and CPs' entry decisions are important issues.[32]

It would be an important research agenda to develop a model that can capture differences between mobile networks and fixed networks. Mobile networks are becoming an increasingly important channel for Internet content delivery. Fixed and mobile Internet networks are inherently different in many dimensions, most importantly in the scarcity of bandwidth for mobile networks imposed by physical laws. These differences are recognized by the recent FCC rule on net neutrality. The new FCC rule, announced on December 21, 2010, reaffirmed the FCC's 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.[33] Our model may lend a new justification for asymmetric

---

[32]Choi and Kim (2010) addresses the dynamic investment issue, but with a monopolistic ISP. Njoroge, Ozdaglar, Stier-Moses, and Weintraub (2010) study investment incentives with multiple ISPs, but neither interconnection between ISPs nor the role of CPs' business models in net neutrality regulation is considered.

[33]See Maxwell and Brenner (2012) and Choi, Jeon, and Kim (2013) for more discussion on the debate of asymmetric regulation and network neutrality.

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 a neutral network, instead of providing somewhat jittery content delivery by serving uniform speed to heterogeneous CPs. If CPs in the mobile networks adopt business models that have more content-usage based charge systems and enable 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 (parameterized by α) for CPs. The analysis can be extended to heterogeneous business models that would be endogenously derived.[34]

---

[34]See Jullien and Sand-Zantman (2012) for a theoretical analysis that endogenizes the choice of business model in the context of congestion pricing and net neutrality. Casadesus-Masanell and Hervàs-Drane (forthcoming) offer a (one-sided market) framework that endogenizes firms' business models where firm profits originate from two revenue sources, the prices directly charged to consumers and advertising revenues.

References

[1] Anderson, Simon, and Stephen Coate. 2005. "Market Provision of Broadcasting: A Welfare Analysis." *Review of Economic Studies* 72 (4): 947-72.

[2] Armstrong, Mark. 1998. "Network Interconnection in Telecommunications." *Economic Journal* 108 (448): 545-64.

[3] Armstrong, Mark. 2006. "Competition in Two-Sided Markets." *Rand Journal of Economics* 37 (3): 668-91.

[4] Armstrong, Mark, and Julian Wright. 2009. "Mobile Call Termination." *Economic Journal* 119 (538): F270-F307.

[5] Bourreau, Marc, Frago Kourandi, and Tommaso Valletti. 2012. "Net Neutrality with Competing Internet Platforms." mimeo.

[6] Casadesus-Masanell, Ramon, and Andrés Hervás-Drane. "Competing with Privacy." *Management Science.* forthcoming.

[7] Choi, Jay Pil. 2006. "Broadcast Competition and Advertising with Free Entry: Subscription vs. Free-to-Air." *Information Economics and Policy* 18 (2): 181-96.

[8] Choi, Jay Pil, and Byung-Cheol Kim. 2010. "Net Neutrality and Investment Incentives." *Rand Journal of Economics* 41 (3): 446-471.

[9] Choi, Jay Pil, Doh-Shin Jeon, and Byung-Cheol Kim. 2013. "Asymmetric Neutrality Regulation and Innovation at the Edges: Fixed vs Mobile Networks." NET Institute Working Paper No. 13-24.

[10] Dessein, Wouter. 2003. "Network Competition in Nonlinear Pricing." *Rand Journal of Economics* 34 (4): 593-611.

[11] Economides, Nicholas, and Benjamin Hermalin. 2012. "The Economics of Network Neutrality." *Rand Journal of Economics* 43 (4): 602-29.

[12] Economides, Nicholas, and Joacim Tåg, 2012. "Net Neutrality on the Internet: A Two-sided Market Analysis." *Information Economics and Policy* 24 (2): 91-104.

[13] Fitzgerald, Drew, and Daisuke Wakabayashi. 2014. "Apple Quietly Builds New Networks." *Wall Street Journal*, February 3.

[14] Hagiu, Andrei, and Robin Lee. 2011. "Exclusivity and Control," *Journal of Economics and Management Strategy* 20 (3): 679-701.

[15] Hermalin, Benjamin, and Michael Katz. 2007. "The Economics of Product-Line Restrictions With an Application to the Network Neutrality Debate." *Information Economics and Policy* 19 (2): 215-48.

[16] Hurkens, Sjaak, and Doh-Shin Jeon. 2012. "Promoting Network Competition by Regulating Termination Charges." *International Journal of Industrial Organization* 30 (6): 541-52.

[17] Jeon, Doh-Sin, Jean-Jacques Laffont, and Jean Tirole. 2004. "On the 'Receiver-Pays' Principle." *Rand Journal of Economics* 35 (1): 85-110.

[18] Johnson, Justin. 2013. "Targeted Advertising and Advertising Avoidance." *Rand Journal of Economics* 44 (1): 128-144.

[19] Jullien, Bruno, and Wilfried Sand-Zantman. 2012. "Congestion Pricing and Net Neutrality." TSE Working paper series 12-327.

[20] Krämer, Jan, Lukas Wiewiorra, and Christof Weinhardt. 2013. "Net Neutrality: A Progress Report." *Telecommunications Policy* 37 (9): 794-813.

[21] Laffont, Jean-Jacques, Scott Marcus, Patrick Rey, and Jean Tirole. 2003. "Internet Interconnection and the Off-Net-Cost Pricing Principle." *Rand Journal of Economics* 34 (2): 370-90.

[22] Laffont, Jean-Jacques, Scott Marcus, Patrick Rey, and Jean Tirole. 2001. "Internet Peering." *American Economic Review* 91(2): 287-91.

[23] Laffont, Jean-Jacques, Patrick Rey, and Jean Tirole. 1998a. "Network Competition I: Overview and Nondiscriminatory Pricing." *Rand Journal of Economics* 29 (1): 1-37.

[24] Laffont, Jean-Jacques, Patrick Rey, and Jean Tirole. 1998b. "Network Competition II: Price Discrimination." *Rand Journal of Economics* 29 (1): 38-56.

[25] Lee, Robin, and Tim Wu. 2009. "Subsidizing Creativity through Network Design: Zero-Pricing and Net Neutrality. " *Journal of Economics Perspective* 23 (3): 61-76.

[26] Maxwell, Whinston, and Daniel Brenner, 2012. "Confronting the FCC Net Neutrality Order with European Regulatory Principles." *Journal of Regulation* 1-2.7.

[27] Milgrom, Paul, and John Roberts. 1990. "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities." *Econometrica* 58 (6): 1255-77.

[28] Njoroge, Paul, Asuman Ozdaglar, Nicolás Stier-Moses, and Gabriel Weintraub. 2010. "Investment in Two Sided Markets and the Net Neutrality Debate." Columbia Business School DRO (Decision, Risk and Operations) Working Paper No. 2010-05.

[29] Schuett, Florian. 2010. "Network Neutrality: A Survey of the Economic Literature." *Review of Network Economics* 9 (2) Article 1.

[30] Vives, Xavier. 2000. *Oligopoly Pricing: Old Ideas and New Tools*, Cambridge, MA: MIT Press.

Appendix A: Proofs

Proof of Lemma 1

Let n_i be the number of consumers subscribing to ISP i in stage 3. For simplicity, let us consider a neutral network and suppose that in stage 1, the ISPs agreed on (q, a). In stage 2, there is Bertrand competition without friction on the CP side and therefore, in equilibrium of this stage, each ISP's price is competitively set equal to the opportunity cost of servicing each CP. This implies that any market share divisions on the CP side yield the same profit to each ISP in equilibrium. In particular, each ISP should receive the same profit whether from capturing the whole CP market or yielding it to the competing ISP. Let p be the equilibrium price and assume that both types buy connections from ISP 1 (for instance) at this price.

Then, ISP 1's total profit depends on (f_1, f_2) and is given by:

$$n_1(f_1, f_2)f_1 + n_1(f_1, f_2)(p - c)q + n_2(f_2, f_1)(p - a - c_o)q.$$

Since ISP 2 loses the CP market, ISP 2's total profit is given by

$$n_2(f_2, f_1)f_2 + n_2(f_2, f_1)(a - c_T)q.$$

Let $(f_1^*(p), f_2^*(p))$ be the equilibrium subscription fees in stage 3. The indifference condition between ISPs' profits in the CP market at stage 2 implies that p satisfies the following condition.

$$n_1(f_1^*, f_2^*)f_1^* + n_1(f_1^*, f_2^*)(p - c)q + n_2(f_2^*, f_1^*)(p - a - c_o)q$$
$$= n_2(f_2^*, f_1^*)f_2^* + n_2(f_2^*, f_1^*)(a - c_T)q.$$

At off-net cost pricing $p = a + c_o$, the equality holds since each ISP i's profit function is the same and is given by

$$n_i(f_i, f_j)f_i + n_i(f_i, f_j)(a - c_T)q.$$

Hence, they will choose $f_1^* = f_2^*$.

When $p \neq a + c_o$, each ISP realizes a different profit from the CP side (both from on-net and off-net traffics) and hence would choose $f_1^* \neq f_2^*$ and the indifference condition is unlikely to hold.

This proves that the off-net pricing induces each ISP to be indifferent among any market share divisions and hence satisfies the local condition for profit maximization.

The same logic can be extended to non-neutral networks and cases in which not all types buy connections.

Proof of Lemma 3

(i) Consider a general demand function $n_i(U - f_i, U - f_j)$ for ISP $i = 1, 2$. Then, ISP i maximizes

$$\Pi_i = n_i(U - f_i, U - f_j)(f_i + \widehat{\pi}^{CP}).$$

From the FOC, in a symmetric equilibrium, we have

$$- m_1(f + \widehat{\pi}^{CP}) + n = 0, \tag{28}$$

where n is the demand per ISP and $m_1 = \partial n_i / \partial(U - f_i)$ and $m_2 = \partial n_i / \partial(U - f_j) < 0$ for $i \neq j$. Note that in a symmetric equilibrium, $f = f(U, \widehat{\pi}^{CP})$.

Equation (28) can be written in more details as

$$-m_1(U - f(U, \widehat{\pi}^{CP}), U - f(U, \widehat{\pi}^{CP}))(f(U, \widehat{\pi}^{CP}) + \widehat{\pi}^{CP}) + n(U - f(U, \widehat{\pi}^{CP}), U - f(U, \widehat{\pi}^{CP})) = 0 \tag{29}$$

Let $a \equiv \partial f / \partial U$ and $b \equiv \partial f / \partial \widehat{\pi}^{CP}$. Totally differentiating (29) leads to

$$-(m_{11} + m_{12})(dU - a dU - b d\widehat{\pi}^{CP})(f(U, \widehat{\pi}^{CP}) + \widehat{\pi}^{CP}) - m_1(a dU + b d\widehat{\pi}^{CP} + d\widehat{\pi}^{CP})$$
$$+ (m_1 + m_2)(dU - a dU - b d\widehat{\pi}^{CP}) = 0 \tag{30}$$

where $m_{11} = \partial^2 n_i / \partial(U - f_i)^2$ and $m_{12} = \partial^2 n_i / \partial(U - f_i)\partial(U - f_j)$. This equation is equivalent to

$$\left[-(1 - a)(m_{11} + m_{12} - m_2)(f + \widehat{\pi}^{CP}) + (1 - 2a)m_1 \right] dU \tag{31}$$
$$+ \left[b(m_{11} + m_{12} - m_2)(f + \widehat{\pi}^{CP}) - (1 + 2b)m_1 \right] d\widehat{\pi}^{CP} = 0$$

Because (31) must hold for any $\left(dU, d\widehat{\pi}^{CP} \right)$, we have two equations:

$$- (1 - a)[(m_{11} + m_{12})(f + \widehat{\pi}^{CP}) - m_2] + (1 - 2a)m_1 = 0; \tag{32}$$

$$b[(m_{11} + m_{12})(f + \widehat{\pi}^{CP}) - m_2] - (1 + 2b)m_1 = 0. \tag{33}$$

Subtracting (33) from (32) leads to

$$(-1 + a - b)[(m_{11} + m_{12})(f + \widehat{\pi}^{CP}) - m_2] + 2m_1(1 - a + b) = 0$$

which is equivalent to

$$(-1 + a - b)[(m_{11} + m_{12})(f + \widehat{\pi}^{CP}) - m_2 - 2m_1] = 0.$$

Using (32), we can simplify the last equality as

$$(-1 + a - b)\frac{m_1}{1 - a} = 0.$$

Therefore, $a - b = \frac{\partial f}{\partial U(\alpha)} + \left|\frac{\partial f}{\partial \widehat{\pi}^{CP}}\right| = 1$ must hold.

(ii) From (33), we have

$$b = -\frac{m_1}{m_1 + (m_1 + m_2) - (m_{11} + m_{12})(f + \widehat{\pi}^{CP})}.$$

Note that $m_1 > 0$ and $m_1 + m_2 > 0$. Under Assumption 3, $(m_{11} + m_{12})(f + \widehat{\pi}^{CP}) \leq 0$ in any symmetric equilibrium with a non-negative profit. This proves $-1 < b < 0$. From $a = 1 + b$, we have $0 < a < 1$.

Proof of Proposition 4 (ii) $(m_{11} + m_{12})(f + \widehat{\pi}^{CP}) \leq 0$ in any symmetric equilibrium with a non-negative profit.

We show that there is no profitable deviation from the off-net cost pricing when the monopoly solution $(q_H^*, q_L^*(\alpha))$ and the associated access charges (a_H^*, a_L^*) are agreed upon. Let $(\overline{q}, \underline{q})$ represent the qualities allocated to high and low types, respectively, in any deviation. 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 the quality $q_L^*(\alpha)$ intended for the low type CPs in the proposed equilibrium, i.e., $(\overline{q}, \underline{q}) = (q_L^*(\alpha), q_L^*(\alpha))$. Since the IC constraint for the high type is binding and high type CPs are indifferent between the two qualities in the monopolistic solution, the best way for ISP i to achieve this deviation is to set a price at an epsilon discount of the off-net cost pricing for $q_L^*(\alpha)$; the price it charges after the deviation is essentially $p_i(q_L^*(\alpha)) = \alpha L u(q_L^*(\alpha))$. The CP side profit (per consumer) from this deviation is given by $\pi_{dev}^{CP} = p_i(q_L^*(\alpha)) - cq_L^*(\alpha) = \alpha L u(q_L^*(\alpha)) - cq_L^*(\alpha)$. Note that a_L^*, the per unit access charge for the low quality delivery the two ISPs agreed on in stage 1, is given by $a_L^* = \alpha L u(q_L^*(\alpha))/q_L^*(\alpha) - c_O$, with the off-net cost pricing of $p_i(q_L^*(\alpha)) = \alpha L u(q_L^*(\alpha)) = (a_L^* + c - c_T)q_L^*(\alpha)$. 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

39

$\pi_i^{CP} = \pi_j^{CP} = \pi^{CP} = (a_L^* - c_T)q_L^*(\alpha)$. This cannot give a higher profit than the upper bound that each ISP can obtain without deviation; otherwise, we have a contradiction because the upper bound is not achieved by the ISPs in the first place.

Consider now the deviation of ISP i to induce both types to consume high quality, i.e., $(\bar{q}, \underline{q}) = (q_H^*, q_H^*)$. This requires ISP i to charge $p_i(q_H^*) = \alpha L u(q_H^*)$ to induce low type CPs to purchase high quality delivery. 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 the CP side is

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

where we use

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

and ISP j's profit from the CP side is

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

Our proof strategy is to show a general result (Lemma 4) that when an ISP attracts all CPs with the same quality of delivery, the ISP's total profit (from the CP and the consumer side) is decreasing with the access charge associated with that quality. We then return to the above specific set-up $(\bar{q}, \underline{q}) = (q_H^*, q_H^*)$. Before proving this, let us describe the setting under which Lemma 4 is obtained.

Specifically, fix $(\bar{q}, \underline{q}) = (q, q)$ and suppose that initially $(\bar{q}, \underline{q}) = (q, q)$ is implemented with the off-net cost pricing such that it generates $\pi_i^{CP} = \pi_j^{CP} = \pi^{CP}$: Then, we get

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

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

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

Then, ISP i's profit from the CP side is

$$N\left[\alpha Lu(q) - s_i cq - (1 - s_i)(c + a' - c_T)q\right]$$

whereas ISP j's profit from the CP side is

$$Ns_j(a' - c_T)q = Ns_j\left[(a - c_T)q + \delta q\right].$$

Note that

$$(c + a' - \delta - c_T)q = \alpha Lu(q).$$

Hence, ISP i's profit from the CP side is

$$N\left[-\delta q + s_i(a' - c_T)q\right] = N\left[-(1 - s_i)\delta q + s_i(a - c_T)q\right].$$

Note that a' will affect (N, s_i, s_j), which is determined in stage 3. Given this, here is the lemma:

Lemma 4. *ISP i's total profit (from the content side and the consumer side) is higher when $\delta = 0$ than when $\delta > 0$.*

Proof. Let $\Pi_i(\delta)$ denote the total profit for ISP i when the access charge is given by $a' = a + \delta$. Then, we have

$$\Pi_i(\delta) = n_i(f_i, f_j)(f_i + \pi^{CP}) - n_j(f_i, f_j)\delta q$$

By using the envelope theorem, we have

$$\frac{d\Pi_i(\delta)}{d\delta} = -n_j q + \left[\frac{\partial n_i}{\partial f_j}(f_i + \pi^{CP}) - \frac{\partial n_j}{\partial f_j}\delta q\right]\frac{df_j}{d\delta}$$

The expression in the square bracket is positive since $\frac{\partial n_i}{\partial f_j} > 0$ and $\frac{\partial n_j}{\partial f_j} < 0$. By totally differentiating the first order conditions for f_i and f_j, we can easily derive a comparative static result that $\frac{df_j}{d\delta} < 0$. The intuition is that an increase in access charge is equivalent to a subsidy by ISP i that captures the whole CP market to ISP j for each consumer ISP j attracts. ISP i competes more aggressively to attract consumers to reduce the subsidy and ISP j also competes more aggressively to attract consumers to increase the subsidy. As a result, competition in the consumer market is intensified. Taken together, we have $\frac{d\Pi_i(\delta)}{d\delta} < 0$, which shows that the winning ISP's overall profit decreases with the access charge. $\square$

As we planned, we now use the above lemma to show that a deviation to induce the quality choice of q_H^* by both types of CPs is not profitable. To see this, consider $(\overline{q}, \underline{q}) = (q_H^*, q_H^*)$, $\delta q_H^* = \alpha\Delta(u(q_H^*) - u(q_L^*(\alpha)))$ and $a' = a_H^*$. Hence, we have

$$
\begin{aligned}
(a - c_T) q_H^* &= (a_H^* - c_T - \delta) q_H^* \\
&= \alpha\left[Hu(q_H^*) - \Delta u(q_L^*(\alpha))\right] - cq_H^* - \alpha\Delta\left[u(q_H^*) - u(q_L^*(\alpha))\right] \\
&= \alpha Lu(q_H^*) - cq_H^*.
\end{aligned}
$$

With the access charge a given by $(a - c_T) q_H^* = \alpha Lu(q_H^*) - cq_H^*$, the off-net cost pricing leads to $(c + a - c_T) q_H^* = \alpha Lu(q_H^*)$. Then, from Lemma 4, we proved that 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 $(c + a - c_T) q_H^* = \alpha Lu(q_H^*)$. Furthermore, the profit in this symmetric equilibrium is what the ISPs could achieve through the off-net cost pricing and should give each ISP a profit smaller than the upper bound. This ends the proof.

In sum, therefore, there is no profitable deviation from the upper bound of the joint profits characterized in Proposition 4.

Proof of Proposition 5 (ii)

Here we show that the upper bound of the joint profits in the neutral network 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 because increasing price for CPs by ISP i attracts no CPs and hence does not affect π_i^{CP} and π_j^{CP}, and decreasing the price only reduces π_i^{CP}.

Second, consider the case in which the 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 $a^* = \alpha Hu(q_H^{FB})/q_H^{FB} - c_O$. Then, off-net cost pricing leads to $p^*(q_H^{FB}) = \alpha Hu(q_H^{FB})$ 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 the 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 $p_i(q_H^{FB}) = \alpha Lu(q_H^{FB})$ and obtain a profit of

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

ISP j's profit is

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

Hence, we can apply Lemma 4. Consider $(\bar{q}, \underline{q}) = (q_H^{FB}, q_H^{FB})$, $\delta q_H^{FB} = \alpha \Delta u(q_H^{FB})$ and $a' = a^*$. As a result, we have

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

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

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

From Lemma 4, the total profit of ISP i upon deviation is smaller than the profit it obtains in a symmetric equilibrium with $(\bar{q}, \underline{q}) = (q_H^{FB}, q_H^{FB})$ and a that satisfies $\left(c + a - c_T\right)q_H^{FB} = \alpha L u(q_H^{FB})$. 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.

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

Appendix B: A Simple Model of Endogenous α

We provide a very simple model to endogenize the CPs' surplus extraction level α based on an informative advertising model [see, for instance, Anderson and Coate (2005) and Choi (2006)]. For simplicity, assume that advertising is the only source of revenues for CPs. To reflect consumers' disutility from being exposed to advertising and CPs' advertising revenues, the total surplus created for CPs and consumers is given by

$$\sum_{\theta} \nu_\theta \theta [1 - \gamma a_\theta + r(a_\theta)] u(q_\theta)$$

$$= \underbrace{\sum_{\theta} \nu_\theta \theta [1 - \gamma a_\theta] u(q_\theta)}_{\text{Consumer Surplus}} + \underbrace{\sum_{\theta} \nu_\theta \theta r(a_\theta) u(q_\theta)}_{\text{CP Surplus}},$$

where a_θ denotes the level of advertising by content type θ with γ being a nuisance parameter and $r(a_\theta)$ represents "normalized" per-consumer advertising revenues for CPs. With this multiplicative form specification, we implicitly assume that the nuisance costs and advertising revenues are proportional to $\theta u(q_\theta)$. We can write that $r(a) = p(a)a$, where $p(a)$ can be considered as the "normalized" inverse demand curve for advertising per consumer. We assume that $r(a)$ is concave in a and achieves its maximum at a^*. This implies that both types of content providers choose the same level of advertising a^* and the CPs' surplus extraction level $\alpha = \frac{r(a^*)}{1-\gamma a^* + r(a^*)}$.

When advertising is the only source of revenues, CPs only care about the advertising revenues and ignore its nuisance value to consumers. Therefore, the optimal advertising level a^* is independent of γ. This implies that α increases in γ.

In this framework, we can also analyze how the changes in advertising technology affect α. For instance, consider the development of targeted advertising technology enabled by the collection of personal information in online marketplaces (see, e.g., Johnson (2013), Casadesus-Masanell and Hervás-Drane (forthcoming)). This development can have effects on both consumers and CPs. Let us assume that targeted advertising is more effective and induces a parallel outward shift in the inverse demand curve for advertising because advertisers are now willing to pay more per advertising. The new inverse demand curve can now be written as $kp(a)$, for $k > 1$. At the same time, let us assume that more personalized and targeted advertising solicits privacy concerns and increases the nuisance cost γ. Then, both effects will induce a higher α.