NET Institute*

www.NETinst.org

Net Neutrality and Investment Incentives

Jay Pil Choi
Michigan State University

Byung-Cheol Kim
Georgia Institute of Technology

Net Neutrality and Investment Incentives[*]

Jay Pil Choi[†]   Byung-Cheol Kim[‡]

October 14, 2008

Abstract

This paper analyzes the effects of net neutrality regulation on investment incentives for Internet service providers (ISPs) and content providers (CPs), and their implications for social welfare. We show that the ISP's decision on the introduction of discrimination across content depends on a potential trade-off between network access fee and the revenue from the trade of the first-priority. Concerning the ISP's investment incentives, we find that capacity expansion affects the sale price of the priority right under the discriminatory regime. Because the relative merit of the first priority, and thus its value, becomes relatively small for higher capacity levels, the ISP's incentive to invest on capacity under a discriminatory network can be smaller than that under a neutral regime where such rent extraction effects do not exist. Contrary to ISPs' claims that net neutrality regulations would have a chilling effect on their incentive to invest, we cannot dismiss the possibility of the opposite.

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

Key Words: Net Neutrality, Investment (Innovation) Incentives, Queuing Theory, Hold-up Problem, Two-sided Markets, Vertical Integration

[†]Department of Economics, Michigan State University, 220A Marshall-Adams Hall, East Lansing, MI 48824 -1038. E-mail: choijay@msu.edu.

[‡]School of Economics, Georgia Institute of Technology, Room 201A Habersham Building, Atlanta, Georgia 30332. E-mail: byung-cheol.kim@econ.gatech.edu.

1 Introduction

This paper analyzes the effects of net neutrality regulation on investment incentives for various players in the Internet market. From its inception, one of the governing principles in the operation of the Internet has been non-discrimination requirements in all relevant performance dimensions, as has been true for traditional telecommunication services such as the telephone network. In 2005, however, the Federal Communications Commission (FCC) changed the classification of Internet transmissions from the category of "telecommunications services" to the category of "information services." As a result, Internet service providers (ISPs) are no longer subject to non-discrimination restrictions. In fact, major telephone and cable operators, which together control about 98 percent of broadband service in the US (as of December 2005),[1] recently expressed an interest to provide multi-tier Internet service, charging content providers (CPs) premium prices for preferential access to the broadband transmission service. In response, a coalition of content providers emerged in an effort to maintain the current status of non-discrimination regime. Their intensive lobbying efforts led to the hot debate – known as the net neutrality debate – in Washington, along with initiatives to legislate a mandate to prevent creating a multi-tier Internet services. Even though the attempt to legislate the net neutrality regulation has failed in Congress for now, the issue is expected to continuously arise in the future.[2]

On October 19, 2007, for instance, the Associated Press (AP) reported that Comcast, the U.S.'s largest cable TV operator and No. 2 Internet provider, interfered with users' access to file sharing sites such as BitTorrent.[3] This practice was an example of discrimination in which ISPs intended to slow down some forms of traffic while giving others priority. Comcast may have had a benign reason for this practice – so called "traffic shaping" – to prevent file-sharing traffic from using up too much bandwidth and affecting the Internet speeds of other subscribers.[4] This interference, however, was certainly a move against the tradition of treating all types of Internet traffic equally – the principle of "net neutrality." Since one person's upload is another's download in file-sharing networks, this type of traffic

---

[1] FCC Form 477 Data.

[2] For detailed explanation and discussion on institutional differences between the EU and U.S. concerning net neutrality regulation, see Chirico, Van de Haar and Larouche (2007), "Network Neutrality in the EU," TILEC, discussion paper, DP 2007-030.

[3] For more detail, see "Comcast Blocks Some Internet Traffic" Oct. 19, 2007, by Peter Svensson, AP.

[4] Peer-to-peer file-sharing applications reportedly account for about 50-90 percent of overall Internet traffic according to a survey in 2007 by ipoque GmbH, a German traffic-management equipment vendor.

management can have a series of repercussions in the network of file sharers. As a result, the incident received nationwide attention and stirred an uproar from users of file-sharing applications who were adversely affected.

To inform this important policy debate, the paper analyzes economic issues associated with net neutrality regulation. Considering that the Internet is a vital medium of communication, information, and commercial activities, maintaining competition and promoting innovation in this market is of paramount importance. Policymakers thus need to act with care and make an informed decision based on rigorous analysis to provide a market environment in which the right investment signals are given when the Internet is involved.

Reflecting the importance of the Internet as a main driver of economic growth and prosperity in the global economy, one of the main issues of the net neutrality debate is the innovation and investment incentive for various parties involved in the market. For instance, ISPs such as Verizon, Comcast, and AT&T oppose network neutrality regulations and claim that such regulations would discourage investment in broadband networks. The logic is that they would have no incentive to invest in network capacity unless content providers who support bandwidth-intensive multimedia Internet traffic pay a premium. In contrast, proponents of network neutrality regulations (comprising mostly consumer rights groups and large Internet content companies such as Google, Yahoo, and eBay) note that the Internet has operated according to the non-discriminatory neutrality principle since its earliest days. They argue that net neutrality has been the main driver of the growth and innovative applications of the Internet. To support their claim, they rely on the so-called end-to-end design principle. Under this design principle, decisions are made "to allow the control and intelligence functions to reside largely with users at the 'edges' of the network, rather than in the core of the network itself."[5] According to them, this creates an environment that does not require users to seek permission from the network owners and thus promotes innovations in Internet applications.

To assess the validity of conflicting claims made by opposing parties, we set up a model that is based on the queuing theory developed in operations research. The reason for this modeling choice is that scarce bandwidth and the potential need for rationing (due to substantial increases in multimedia usage of the Internet) are the root causes of the debate. The queuing theory literature has shown that assuming a constant Poisson arrival rate of

---

[5]Cerf (2006).

content requested by each consumer generates a process that is a good approximation of congestion in computer networks.

With the adoption of such microfoundations in a setup with a monopolistic network operator and two application providers, we provide a formal economic analysis on the effects of net neutrality regulation on investment incentives for Internet service providers (ISPs) and content providers (CPs), and their implications for social welfare. More specifically, we first compare the market equilibrium in which the monopolistic ISP is allowed to provide a two-tiered service by selling the "fast-lane" to only one content provider to the equilibrium in which it cannot discriminate the delivery speed of content. This comparison of short-run equilibrium yields two major findings. First, both content providers may engage in a Prisoners' dilemma type of game to receive the first priority in the delivery of content and be worse off in a discriminatory network. The ISP's decision of whether or not it will prefer the discriminatory regime to the neutral network depends on a potential trade-off between its network access fee from end users and the revenue from CPs through the trade of the first-priority. Second, the short-run effect of net neutrality regulation on social welfare depends on the relative magnitudes of content providers' cost/quality asymmetry and the degree of content differentiation. In particular, we show that social welfare is higher under net neutrality if the asymmetry across content providers is sufficiently small.

Additionally and more importantly, we study the long-run effects of net neutrality regulation on the ISPs' investment incentives. We find that there are two channels through which net neutrality regulation can have impacts on the ISPs' investment incentives: the *network access fee effect* and the *rent extraction effect*. In the network with net neutrality, capacity expansion speeds up the delivery of content uniformly, thereby enabling the ISP to charge more for access. Similarly, in the discriminatory network, capacity expansion also increases the delivery speed of content and thus allows the ISP to charge a higher network fee. However, because such effect occurs asymmetrically across different priority classes, we cannot tell unambiguously under which regime the effect of capacity extension is larger. Capacity expansion also affects the sale price of the priority right under the discriminatory regime. Because the relative merit of the first priority, and thus its value, becomes relatively small for higher capacity levels, the ISP's incentive to invest on capacity under a discriminatory network is smaller than that under a neutral regime where such rent extraction effects do not exist. As a result, the ISP's investment incentive hinges upon the relative

4

magnitudes of these two potentially opposing effects. Once again, it is *a priori* ambiguous whether the ISP has greater incentive to invest in capacity in a neutral network or a discriminatory one. Contrary to ISPs' claims that net neutrality regulations would have a chilling effect on their incentive to invest, we cannot dismiss the possibility of the opposite.

We also study the effects of net neutrality regulation on application/content providers' incentives to invest in cost reduction/quality enhancement. Because the monopolistic ISP can expropriate some of the investment benefits made by content providers through the trade of first-priority delivery in a discriminatory network, content providers' investment incentives can be higher under the net neutrality regime. This implies that the ISP's payoff is not necessarily increasing in its ability to extract rents from CPs when the adverse effects on CPs' investment incentives are taken into account. As a result, the ISP may wish to limit its ability to extract rent, if such a commitment mechanism is available, to mitigate the countervailing dynamic effect on innovation incentives for CPs.

We thus find that the relationship between net neutrality regulation and investment incentives for network operators and application/content providers is subtle, and it is not easy to draw general unambiguous conclusions. However, our model informs policymakers and regulators by identifying important effects that are expected both in the short run and long run and showing the mechanism through which such effects interact.

The remainder of the paper is organized in the following way. The next section offers a brief literature review of papers addressing net neutrality issues. Section 3 sets up a preliminary model of network markets to analyze the effects of net neutrality regulation on competition and social welfare. Sections 4 and 5 analyze the effects of net neutrality regulations on investment incentives of ISPs and CPs, respectively. In section 6, we provide a brief analysis with discussion about various issues around the debate of net neutrality such as heterogeneity in delay costs across content, quality degradation of information packets, and vertical integration between ISP and CP. Section 7 closes this article with concluding remarks along with suggestions for further possible extensions of our basic analysis. Most of the proofs for lemmas and propositions are relegated to the Appendix.

2 Related Literature

Net neutrality regulations have been a hotly debated topic discussed with passion by both proponents and opponents alike. The discussion so far, however, has been rich in rhetoric but short on rigorous economic analysis. There are several notable exceptions.[6]

Hermalin and Katz (2007) consider a situation in which ISPs serve as a platform to connect content providers with end consumers. As in our paper, they adopt a framework of the so-called two-sided markets to analyze the effects of net neutrality regulation. More specifically, they consider heterogeneous content providers whose products are vertically differentiated. Without any restrictions, ISPs can potentially offer a continuum of vertically differentiated services to heterogeneous content providers. They formally model the network neutrality regulations as product line restrictions that require ISPs to provide only one service level (a single tier of Internet service). To analyze the effects of regulation, they compare the single-service level equilibrium to the multi-service level equilibrium. They show that net neutrality regulation has the following effects. Content providers who would otherwise have purchased a low-quality service are excluded from the market. That is, content providers at the bottom of the market – the ones that a single-product restriction is typically intended to aid – are almost always harmed by the restriction. Content providers in the "middle" of the market utilize more efficient and higher quality service, which favors the net neutrality regulation. Content providers at the top of the market utilize less efficient and lower quality service than the one that would have been used in the absence of regulation, which obviously favors the discriminatory network. The overall welfare effect of such regulation can be ambiguous, but they argue that the effects are often negative. The analysis of Hermalin and Katz, however, does not consider the congestion effect in the provision of Internet service. More importantly, their analysis is static in the sense that they do not investigate investment incentives of content providers and ISPs, the central concern in the net neutrality debate. Therefore, our research thus complements that of Hermalin and Katz (2007).

In terms of the policy questions asked as well as basic framework, our research is closest to Cheng et al. (2006), who develop a game-theoretic model of competition between two

---

[6]See also Economides (2007) and Kocsis and de Bijl (2007). In addition, there is an extensive discussion on net neutrality by lawyers. See, for instance, Wu (2003), Yoo (2006), and van Schewick (2007) and references cited therein.

content providers in a Hotelling framework. They investigate the effects of net neutrality regulation on ISPs' incentives to expand capacity in addition to addressing the question of who gains and who loses as a result of regulation. However, there are several conspicuous differences between our paper and theirs. In this study, we intend to go one step further by analyzing the effects of the regulation on content providers' incentives to provide innovative services. We find that the hold-up problem can prevail under a discriminatory regime and thus *ex ante* the ISP might prefer to commit to the maintenance of a neutral network. In addition, we find somewhat different results from Cheng et al. even with a similar framework. For instance, we find that it is not easy to draw general clear-cut conclusions about the relationship between net neutrality regulation and innovation incentives of either ISPs or CPs. In contrast, they find that if the principle of net neutrality is abandoned, the broadband service provider definitely stands to gain from the arrangement, as a result of extracting the preferential access fees from the content providers. Another example is that they find that the ISP's incentive to expand its capacity is unambiguously higher under net neutrality, while we find such an outcome is just one possibility. Finally, the analysis of Cheng et al. (2006) lacks analytical consistency in the formulation of waiting time under non-neutrality and employs somewhat *ad hoc* assumptions in the analysis of capacity expansion incentives. This paper has eliminated such problems.

Economides and Tåg (2007) provide an economic analysis on net neutrality in a two-sided market framework. The main focus in their article differs from ours. They are particularly interested in the effects of net neutrality regulation on pricing schemes on both sides of the market and on social welfare in the short run. In this paper, we discuss the effects of neutrality regulation on the players' dynamic innovation incentives. Thus, our research strongly complements theirs.

Finally, Valletti and Cambini (2005) analyze the network operators' incentives to invest in networks with different quality levels, as in our paper. They show that quality has an impact on all calls initiated by customers (destined both on-net and off-net) and "tacit collusion" takes place even in a symmetric model with two-part pricing because firms tend to underinvest in quality. However, their focus is on the impact of *two-way access charges* on the investment incentives in communication networks that require interconnection for off-net traffic whereas our analysis concerns the impacts of net neutrality regulation on investment incentives of a network operator that serves as a platform for two-sided markets.

3 A Model of Net Neutrality

We consider a situation in which online content providers deliver their contents to end consumers through a broadband network that is provided by a monopolistic Internet service provider (ISP). For instance, we can envision a specific geographic market in which Comcast is a monopolistic ISP and content providers such as Yahoo and Google deliver their contents at the end users' requests.[7] There is no universally accepted definition of net neutrality. For the sake of analysis, in this paper we simply define net neutrality as non-discrimination in the delivery of content (packets) through the network.[8]

3.1 The Basic Model

The monopolistic ISP sells its network connection to end users at price a. There are two content providers who compete to deliver content to end users. Under net neutrality, the ISP cannot discriminate between content providers in the delivery speed of contents. For simplicity, let us assume that under net neutrality the ISP provides content providers with the network line at no charge.[9] In contrast, without net neutrality regulation, preferential treatment for a particular content provider is no longer prohibited. Then, the ISP can sell the first-priority, the right to be served ahead of the other, to either one of the two content providers. As will be explained in further detail, we adopt a general framework that can capture various manners in which the first priority can be sold. In this sense, we consider access-tiering as a practice that violates network neutrality, instead of port blocking or quality degradation.[10]

As in standard queuing models, we assume that the arrival rate of each consumer follows a Poisson distribution with λ. The processing times of all jobs in the network are exponentially distributed with the same mean $1/\mu$, where the service rate μ is determined by network capacity. This setup is well-known to be a very good approximation for the arrival process in real systems, in which the number of customers is sufficiently large so

---

[7]Our model complements Kocsis and de Bijl (2007) that consider the situation where there is effective competition between a small number of network operators. We mention a possible extension to such a direction in section 7.

[8]In other words, we use net neutrality and non-discrimination interchangeably throughout this article. However, see Wu (2003) who considers net neutrality as an end and non-discrimination as a mean toward that end.

[9]To quote AT&T CEO Edward Whitacre, content providers "use my lines for free." See "Rewired and Ready for Combat," *Business Week Online*, November 7, 2005.

[10]See Kocsis and de Bijl (2007) for these types of violations of net neutrality.

that the impact of a single customer on the performance of the system is very small, and all customers' decisions to use the system are independent of other users'. In the short-run analysis, capacity μ is assumed to be fixed. In the long-run analysis in which investment incentives are investigated, capacity μ is endogenously derived. In the neutral network regime, each packet is treated equally and delivered on a first-come, first-served basis. In the discriminatory network regime, packets with priority class are delivered first, ahead of any other packets.

Consumers, whose mass is normalized to one, are heterogenous with respect to their preferences toward two content services in the Hotelling manner. By setting CP1 and CP2 to be located at the left and right ends of a line segment whose length is normalized to one, a consumer located at x pays the transport cost of tx and $t(1-x)$ to consume CP1's and CP2's services, respectively. As usual, the transport cost per unit distance, t, can represent the degree of product differentiation. As in Mendelson (1985), we assume that consumers whose rate of content request is given by λ derive a gross utility of $v(=V(\lambda))$ from either content service, and this reservation value of content service is assumed to be sufficiently high to ensure that the market is fully covered both in the neutral and discriminatory networks.[11]

As in Cheng et al. (2006), Choi (2006), and Economides and Tåg (2007), we assume that content providers adopt a business model that offers their services without any direct charge, but generate their revenues through advertisement. Advertisement revenues depend on their market shares. More specifically, each content provider i earns a revenue stream r from advertisers for each consumer's content request ("click-throughs") it serves. The cost of serving each consumer's request is given by c_i, where $0 \leq c_1 \leq c_2$ without loss of generality.[12] Thus, content provider i's mark-up per each consumer's click-through and the corresponding profit per consumer are respectively given by $(r - c_i)$ and $(r - c_i)\lambda q_i$ where q_i denotes the market share for content provider i.[13]

The sequence of the players' choices are as follows. In the discriminatory network regime,

---

[11] Here we treat the demand parameter λ as exogeneous. However, it can depend on the delivery speed of content to end users in a more general model. For instance, it is possible that end users may abort content requests in the face of long delays and leave the queue.

[12] Alternatively, we could introduce asymmetry in the revenue stream parameter r instead of the cost parameter with the same qualitative results.

[13] Abstracting from direct payments between content providers and end users simplifes the analysis considerably. The exploration for the implications of direct payment will be an important extension of this basic model as explained in section 7.

the ISP can first sell the priority service through a trading process to one content provider; in the neutral network this stage does not apply. Second, the ISP posts a network access fee, a, to end users. Given the allocation of priority classes and the network subscription fee, end users choose one of the content providers. As usual, the analysis for this game proceeds by using backward induction, and the equilibrium concept employed here is that of subgame perfect Nash equilibrium.

3.2 Preliminaries: Congestion in the M/M/1 Queuing System

To model congestion in the network, we adopt the standard framework of the M/M/1 queuing system that has been widely used by many scholars in operations research to study congestion problems and priority pricing (See Naor, 1969; Balanchandran, 1972; Edelson and Hilderbrand, 1975; Mendelson and Whang, 1990).[14] The reason for this modeling choice is that scarce bandwidth and the potential need for rationing (due to substantial increases in multimedia usage of the Internet) are the root causes of the debate. This micro-foundation yields nice properties with which we can work for our analysis without any *ad hoc* assumptions.

In a neutral network where all packets are treated equally without any priority classes, it is a standard result in the queuing theory that each consumer has the expected waiting time of

$$w = \frac{1}{\mu - \lambda} \tag{1}$$

where λ denotes the gross arrival rate at the network (with the normalization of consumer mass to one) and μ is the capacity of the network with $\mu > \lambda$. As is intuitively expected, the waiting time increases in λ, but decreases in μ. If we normalize the delay cost per unit time to one, then the expression for the waiting time equals that of the waiting cost.[15]

On the other hand, in the discriminatory network with two priority classes, consumers' waiting costs depend on the priority classes to which their packets are designated. In the

---

[14]See Gross and Harris (1998) for a standard reference on the queueing theory.

[15]In the basic model, we assume that all content has the same delay cost per unit time. This assumption can be relaxed by assuming heterogeneity in delay costs across content and applications. See Section 6.

non-preemptive discriminatory network,[16] Gross and Harris (1998, pp.146-147) show that a consumer who requests content designated to the first-priority class has an expected waiting time of

$$w_1 = \frac{1}{\mu - \lambda_1} \tag{2}$$

where λ_1 is the total amount of traffic from consumers who request the content with first-priority.[17] In contrast, the consumer who requests content without first priority faces the expected waiting time of

$$w_2 = \frac{\mu}{\mu - \lambda} w_1 = \frac{\mu}{\mu - \lambda} \frac{1}{\mu - \lambda_1} \tag{3}$$

Based on these standard results in the queuing theory for the M/M/1 system, we can derive intuitive results that play important roles in the subsequent analysis. First, in a discriminatory network, a consumer experiences a longer delay by subscribing to the basic service instead of the premium one, i.e.,

$$\text{Fact 1.} \quad w_2 > w > w_1 \quad \text{for} \quad \mu > \lambda.$$

We can easily establish this fact by examining the ratio w_2 to w_1, i.e., $w_2/w_1 = \mu/(\mu-\lambda) > 1$. As a related fact, we note that the relative ratio of w_2 to w_1 is a constant, regardless of the distribution of the total traffic across different priority classes.

In addition, by taking the first derivative of waiting cost differential across classes of services with respect to the network capacity, we find that the quality difference measured in waiting costs becomes smaller as the network capacity increases, i.e.,

$$\text{Fact 2.} \quad \frac{\partial}{\partial \mu}(w_2 - w_1) < 0.^{[18]}$$

This is because the marginal saving in waiting time for the fast-lane from capacity expansion decreases as the capacity level becomes high. It is noteworthy at this stage that the above Fact 2 will play a crucial role in some of the findings concerning the ISP's incentive to invest

---

[16] In discriminatory networks, there are two possible priority schemes: preemptive and non-preemptive schemes. In the preemptive scheme, the customer request with the priority is allowed to be serviced immediately, even if another without priority is already present in service. In the non-preemptive scheme, the customer request with the priority simply goes to the head of the queue to wait its turn without interrupting the service of a customer request already in process.

[17] Following convention in queuing theories, the smaller number represents the higher priority.

[18] $\frac{\partial}{\partial \mu}(w_2 - w_1) = -\frac{\lambda}{(\mu-\lambda_1)(\mu-\lambda)}\left(\frac{1}{\mu-\lambda_1} + \frac{1}{\mu-\lambda}\right) < 0.$

in network capacity.

4 Net Neutrality and ISPs' Investment Incentives

In this section, we study the effects of net neutrality regulations on ISPs' investment incentives. As usual, we apply backward induction to analyze the investment incentives. We first analyze short-run equilibrium in the network market given network capacity μ. Then, we extend the analysis to incorporate dynamic considerations since the net neutrality debate centers around future investment and innovations,[19] noticing that one of the main issues in the debate is how the broadband operator's incentive to expand capacity in infrastructure would be affected by allowing preferential transmission of content. We address this long-run issue by investigating the ISP's marginal change in its profit with respect to the capacity parameter μ.

4.1 Short-Run Analysis with a Fixed Level of Capacity

4.1.1 Equilibrium in a Neutral Network: A Benchmark Case

With the net neutrality regulation, there are no priority classes in content delivery: each packet is treated equally on the basis of first-come, first-served. Each end user chooses one of the two content providers, CP1 and CP2, that provides higher net surplus. In the Hotelling model of end users, the marginal consumer x^* who is indifferent between two content providers in a neutral network is defined as[20]

$$v - \frac{1}{\mu - \lambda} - tx^* - a = v - \frac{1}{\mu - \lambda} - t(1 - x^*) - a, \tag{4}$$

where consumers whose preferences are represented by $x < x^*$ choose CP1 and those with $x > x^*$ choose CP2. With two symmetrically positioned content providers, the market for content provision is equally split between the two firms with each content provider serving half of the market, i.e., $x^* = 1/2$. We assume that v is sufficiently large so that it is in the best interest of the monopolistic ISP to serve all end users.

---

[19]Wu (2003), for instance, states that "[t]he arugment for network neutrality must be understood as a concrete expression of a system of belief about innovation (p. 145)."

[20]The following equality is based on the assumption that there is no direct payment from end users to content providers, which simplifies the analysis.

The ISP's profit maximization problem is thus given by

$$\max_{a} \pi_m = a \quad \text{s.t.} \quad v - \frac{1}{\mu - \lambda} - tx^* - a \geq 0, \tag{5}$$

where the constraint is needed to ensure that the market is covered. Then, we can derive the equilibrium network subscription fee and each content provider's profit as

$$\pi_m^* = a^* = v - \frac{1}{\mu - \lambda} - \frac{t}{2}; \qquad \pi_i^* = \frac{r - c_i}{2}\lambda \quad \text{for } i = 1, 2. \tag{6}$$

4.1.2 Equilibrium in a Discriminatory Network

If the ISP is allowed to charge content providers for the higher priority class, consumers will face different expected waiting times according to their choices of content services. Let us assume that the low-cost content provider, CP1, obtains the first-priority. This means that CP1's content is entitled to be served ahead of CP2's.[21] Then, the consumer at $\widetilde{x}$, who is indifferent between the premium service provided by CP1 and the basic service provided by CP2, is characterized by the equality of the net surpluses from each choice:[22]

$$v - \frac{1}{\mu - \widetilde{x}\lambda} - t\widetilde{x} - a = v - \frac{\mu}{\mu - \lambda}\frac{1}{\mu - \widetilde{x}\lambda} - t(1 - \widetilde{x}) - a. \tag{7}$$

The waiting costs are based on the M/M/1 queuing system with two priority classes and no preemption. Note that, unlike Cheng et al. (2006), a consumer's waiting cost for content without first-priority is adversely affected by the volume of priority traffic. In particular, the consumer who requests CP2's content faces a higher waiting cost than that in the neutral network.

By comparing (4) and (7), we can derive an intuitive result that the content provider with first-priority has a larger market share than the one without it, i.e., $\widetilde{x} \geq x^* = 1/2$ due to the difference in waiting times. The consumer located just to the right of $x^* = 1/2$ receives a discretely higher utility by choosing the content delivered at the premium rate, but faces a marginally higher transportation cost. More consumers will keep choosing CP1's content until the waiting cost saved by this choice is equal to the increased disutility from the choice of lower priority content. This process may lead to a corner solution. To

---

[21]Later we demonstrate that the low-cost firm receives the priority as an equilibrium outcome.
[22]We use a tilde to denote variables associated with a discriminatory regime.

see this, note that

$$\frac{\partial w_2}{\partial \lambda_1} = \frac{\mu}{\mu - \lambda} \frac{1}{(\mu - \lambda_1)^2} > \frac{1}{(\mu - \lambda_1)^2} = \frac{\partial w_1}{\partial \lambda_1} > 0$$

The condition above states that as more consumers subscribe to the CP with the first priority, the waiting costs for both types of CPs increase, but the marginal effect on the waiting cost for non-priority CP is greater. As a result, we may end up a situation in which all consumers subscribe to the CP with the first priority.[23] To prevent this outcome from prevailing, we need that the two CPs are sufficiently differentiated. More specifically, for the adjustment process to yield a stable interior equilibrium and the "right" signs for comparative statics results, we assume the following condition:

$$
\begin{aligned}
\Gamma_1'(\widetilde{x}) \quad &> \quad \Gamma_2'(\widetilde{x}) \text{ for all } \widetilde{x} \in [1/2, 1], && (8)\\
\text{where } \Gamma_1(\widetilde{x}) \quad &= \quad \frac{1}{\mu - \widetilde{x}\lambda} + t\widetilde{x} \text{ and } \Gamma_2(\widetilde{x}) = \frac{\mu}{\mu - \lambda} \frac{1}{\mu - \widetilde{x}\lambda} + t(1 - \widetilde{x}).
\end{aligned}
$$

By taking the derivatives of $\Gamma_1(\widetilde{x})$ and $\Gamma_2(\widetilde{x})$, we can explicitly write the condition (8) as

$$\frac{\lambda^2}{(\mu - \lambda)(\mu - \widetilde{x}\lambda)^2} < 2t \qquad (9)$$

The following lemma specifies a sufficient condition for (9) to hold.

Lemma 1 If $\mu > \frac{3\lambda}{2}$, then the stability condition holds with $\frac{\lambda^2}{(\mu-\lambda)(\mu-\widetilde{x}\lambda)^2} < 2t$.

Proof. Note that $\frac{\lambda^2}{(\mu-\lambda)(\mu-\widetilde{x}\lambda)^2} = \frac{\lambda}{(\mu-\lambda)(\mu-\widetilde{x}\lambda)} \frac{\lambda}{(\mu-\widetilde{x}\lambda)} = \frac{\lambda}{(\mu-\widetilde{x}\lambda)}(2\widetilde{x} - 1)t$. The last equality comes from equation (7) that defines $\widetilde{x}$. Thus, the stability condition holds if $\frac{\lambda}{(\mu-\widetilde{x}\lambda)}(2\widetilde{x} - 1) < 2$. Notice that the LHS of the inequality above is increasing in $\widetilde{x}$ whose maximum value can be 1. It can easily be seen that if $\mu > \frac{3}{2}\lambda$, the above inequality is satisfied even for $\widetilde{x} = 1$.[24] ∎

In the rest of the paper, we assume that $\mu > \frac{3\lambda}{2}$ to focus on the stable equilibrium. Under this maintained assumption, the following comparative statics result shows that the

---

[23]Ironically, in this outcome no one has priority because everyone is treated equally within the priority class.

[24]Alternatively, we can assume that the transportation cost parameter is sufficiently high that the critical consumer's location in a discriminatory network is located between 1/2 and 3/4 for the relevant parameter values. Then, it can be shown that the condition holds.

market share of the CP with the priority for content delivery decreases as the ISP's capacity increases. The main intuition for this result is that an increased capacity of ISP makes congestion less important and reduces the relative quality differential (i.e., waiting costs) across the two CPs.

Lemma 2 $\frac{d\tilde{x}}{d\mu} < 0$.

Proof. By totally differentiating (7), we find the following relationship of

$$sign\left(\frac{d\tilde{x}}{d\mu}\right) = sign\left(\frac{\lambda^2}{(\mu - \lambda)(\mu - \tilde{x}\lambda)^2} - 2t\right). \tag{10}$$

By lemma (1), $\frac{\lambda^2}{(\mu-\lambda)(\mu-\tilde{x}\lambda)^2} - 2t < 0$ under the maintained assumption of $\mu > \frac{3\lambda}{2}$. Therefore, $\frac{d\tilde{x}}{d\mu} < 0$.[25] ∎

In the discriminatory network, the ISP's profit is given by

$$\max_{a} \tilde{\pi}_m = a + f \qquad \text{s.t} \qquad v - \frac{1}{\mu - \tilde{x}\lambda} - t\tilde{x} - a \geq 0, \tag{11}$$

where f denotes the ISP's revenue from the provision of first-priority to CP1. We do not specify a particular trading mechanism that determines f. Instead, we take a more general approach that can encompass various trading protocols. When both CPs compete to acquire the priority right, the winner is typically determined by the maximum willingness to pay. Note that each content provider knows that its market share will be $\tilde{x}$ if it acquires the priority right and $(1 - \tilde{x})$ if the other CP acquires the priority. Consequently, each content provider's maximum willingness to pay for the priority service is given by $(r - c_i)(2\tilde{x} - 1)\lambda$. For instance, if the priority right is sold through a first price ascending auction, CP1 will receive the priority at the price of $f = (r - c_2)(2\tilde{x} - 1)\lambda$, which is the CP2's maximum willingness to pay for the right.[26] Alternatively, we can also envision a situation in which the ISP makes sequential take-it-or-leave-it offers: the ISP makes the first offer to CP1 and if it is not accepted by CP1 and it will make another offer to CP2. In such a scenario, the ISP can extract all surplus from CP1 by charging $f = (r - c_1)(2\tilde{x} - 1)\lambda$. We adopt

---

[25]In fact, $\mu > \frac{3\lambda}{2}$ is the necessary condition under which $\frac{d\tilde{x}}{d\mu}$ is defined as a real number.

[26]Economides (2007) discusses several consequences of the departure from net neutrality regulation based on the auction of prioritization through which only one group of content providers is entitled to the right to the fast lane.

a framework that can encompass both scenarios above and the full range between them that represents different surplus divisions between the ISP and the CP that acquires the priority.

More specifically, let θ $(0 \leq \theta \leq 1)$ denote the ISP's bargaining power in that it measures the proportion of rent extraction from the low-cost content provider, CP1.[27] The price of the first priority is given by

$$
\begin{aligned}
f|_{\theta \in [0,1]} &= \theta(r - c_1)(2\widetilde{x} - 1)\lambda + (1 - \theta)(r - c_2)(2\widetilde{x} - 1)\lambda \qquad (12) \\
&= [r - \theta c_1 - (1 - \theta)c_2](2\widetilde{x} - 1)\lambda
\end{aligned}
$$

For instance, the case where the ISP is able to extract the entire rent from the low-cost content provider by making sequential take-it-or-leave-it offers is characterized by $\theta = 1$ with $f|_{\theta=1} = (r - c_1)(2\widetilde{x} - 1)\lambda$. The other case where the right to the priority is traded through the first-price bid auction scheme is captured by the special case of $\theta = 0$ with $f|_{\theta=0} = (r - c_2)(2\widetilde{x} - 1)\lambda$. All the intermediate cases are captured by some $\theta \in (0, 1)$. As expected, the more bargaining power the ISP has, the higher the priority price will be, which is easily shown as $\frac{\partial f}{\partial \theta} = (c_2 - c_1)(2\widetilde{x} - 1)\lambda \geq 0$.

The ISP's profit in a discriminatory network thus is given by

$$
\widetilde{\pi}_m^* = \left(v - \frac{1}{\mu - \widetilde{x}\lambda} - t\widetilde{x}\right) + [r - \theta c_1 - (1 - \theta)c_2](2\widetilde{x} - 1)\lambda. \qquad (13)
$$

When the ISP assigns the right to the fast lane to the low-cost content provider at the price in (12), each content provider's profit is respectively given by

$$
\begin{aligned}
\widetilde{\pi}_1^* &= (r - c_1)\widetilde{x}\lambda - [r - \theta c_1 - (1 - \theta)c_2](2\widetilde{x} - 1)\lambda \qquad (14) \\
\widetilde{\pi}_2^* &= (r - c_2)(1 - \widetilde{x})\lambda
\end{aligned}
$$

4.1.3 The Short-Run Effects of Net Neutrality on Players

We now analyze the effects of net neutrality regulation on various players. For instance, the effects of regulation on the ISP's profits can be analyzed by comparison of (6) and (13).

---

[27] We does not pin down detailed microfoundations for the bargaining process, because such an issue is not the focus of our paper.

We find the following potential trade-off: without net neutrality the ISP earns less profit from consumers due to the decreased network access fee (a), but gains from trading the priority to the low-cost content provider(f).

Lemma 3 *The network access fee in a discriminatory network is lower than that in a neutral network, i.e., $\widetilde{a} < a^*$.*

Proof. Note that $a^* = v - \frac{1}{\mu-\lambda} - \frac{1}{2}t$ and $\widetilde{a} = v - \frac{1}{\mu-\widetilde{x}\lambda} - t\widetilde{x}$. The difference in network access fee is given by

$$a^* - \widetilde{a} = \frac{1}{\mu-\widetilde{x}\lambda} - \frac{1}{\mu-\lambda} + t(\widetilde{x} - \frac{1}{2}). \tag{15}$$

Recalling that $\widetilde{x}$ is defined by (7), $\widetilde{x}$ satisfies the equality of

$$t(2\widetilde{x} - 1) = \frac{1}{\mu-\widetilde{x}\lambda}\frac{\mu}{\mu-\lambda}. \tag{16}$$

Therefore, by dividing (16) by two, then substituting (16) into (15),

$$
\begin{aligned}
a^* - \widetilde{a} &= \frac{1}{\mu-\widetilde{x}\lambda} - \frac{1}{\mu-\lambda} + \frac{1}{2}\frac{1}{\mu-\widetilde{x}\lambda}\frac{\mu}{\mu-\lambda} \\
&= \frac{(2\widetilde{x}-1)\lambda}{2\left(\mu-\widetilde{x}\lambda\right)\left(\mu-\lambda\right)} > 0 \quad \text{because } \widetilde{x} > \frac{1}{2}.
\end{aligned}
$$

∎

In the absence of regulation, the ISP will choose to introduce the premium service when its gain from prioritization is sufficiently high. Proposition 1 summarizes the effects of introducing two-tiered services on all parties concerned.

Proposition 1 *(a) $\widetilde{\pi}_m^* \gtreqless \pi_m^*$ iff $r \gtreqless \bar{r}$ where $\bar{r} \equiv \theta c_1 + (1-\theta)c_2 + \frac{t}{2\lambda} + \frac{1}{(2\widetilde{x}-1)\lambda}\left(\frac{1}{\mu-\widetilde{x}\lambda} - \frac{1}{\mu-\lambda}\right)$;*
(b) $\widetilde{\pi}_1^* > \pi_1^*$ iff $r - c_2 < (c_2 - c_1)(1 - 2\theta)$;
(c) $\widetilde{\pi}_2^* \leq \pi_2^*$ for $\forall\ r, c_i, \theta, \lambda$; and
(d) Aggregate consumer welfare increases.

Proof. The statements in (a), (b), and (c) can be proved in a straightfoward manner by comparing the expressions for profits across the regimes. Concerning the statement in (d), let us denote the aggregate consumer welfare with the neutral network and the discriminatory network by CS and $\widetilde{CS}$, respectively. Notice that the marginal consumers in the neutral network and the discriminatory network are located at $x^* = 1/2$ and $\widetilde{x}(>$

1/2), respectively, and they receive zero payoffs. This implies that $CS = 2\int_0^{1/2} txdx$ and $\widetilde{CS} = \int_0^{\widetilde{x}} txdx + \int_0^{1-\widetilde{x}} txdx$. Therefore, $\widetilde{CS} - CS = \int_{1/2}^{\widetilde{x}} txdx - \int_{1-\widetilde{x}}^{1/2} txdx > 0$. ∎

Proposition 1 identifies the beneficiaries and losers of net neutrality regulation. Part (a) states that the ISP's profit is higher with a discriminatory network if the advertising revenue from consumers' click-throughs (r) is sufficiently high. In such a case, market share is more important and CPs compete more aggressively to obtain the first priority in a discriminatory network. As a result, the ISP receives a higher price for the premium service, which can outweigh any potential loss in access fees from end users. This also implies that unless r is sufficiently high, the ISP will endogenously choose the equal treatment of both content providers even though the net neutrality is *not* required. Parts (b) and (c) concern the comparison of the CPs' payoffs under different regimes. The low-cost content provider who obtains the first priority can have a higher payoff in the discriminatory regime if the cost differential between the two content providers is sufficiently large. In contrast, the high-cost content provider is always worse off from the introduction of priority classes. They also show the possibility that both content providers may engage in a Prisoners' dilemma type of game to receive the first priority in the delivery of content in the sense that they end up with lower payoffs, whereas the ISP prefers a discriminatory network. This case takes place if $r > \max[\overline{r}, c_2 + (c_2 - c_1)(1 - 2\theta)]$.

4.1.4 The Effects of Net Neutrality on Short-Run Social Welfare

With the Hotelling model for the end users, social welfare analysis of two-tiered services is fairly straightforward: there is no demand effect with pricing, as long as the market is covered. However, there are three types of costs we need to compare to analyze the effects of two-tiered pricing on social welfare: i) total service costs, ii) total transportation costs and iii) total delay costs. The following series of lemmas respectively examine the effects of these factors on the short-run social welfare.

First, the discriminatory regime allows the low-cost content provider to expand its market share through speedier delivery of its content. As a result, the efficiency in terms of production cost minimization favors the discriminatory network. We can easily calculate the cost saved under a discriminatory regime by calculating the difference in total service costs between two distinct regimes:

Lemma 4 *Total service cost under neutrality regime is higher than that under discriminatory regime.*

Proof. *Let S and $\widetilde{S}$ denote the total service cost in a neutral network and in a discriminatory network, respectively.*

$$\begin{aligned}
\Delta S &\equiv& S - \widetilde{S} = \frac{c_1 + c_2}{2}\lambda - (\widetilde{x}c_2 + (1 - \widetilde{x})c_1)\,\lambda \\
&=& \left(\widetilde{x} - \frac{1}{2}\right)(c_2 - c_1)\lambda \geq 0 \ since \ \widetilde{x} > 1/2 \ and \ c_1 \leq c_2.
\end{aligned}$$

∎

Second, recalling that the total transportation costs are minimized when the critical consumer is located at the mid-point, the two-tiered pricing with $\widetilde{x} > 1/2$ is inefficient in terms of transportation cost minimization. We can easily check that the transportation cost in the discriminatory network is higher than that in the neutral network.

Lemma 5 *The transportation cost in the discriminatory network is higher than that in the neutral network:*

Proof. *Let T and $\widetilde{T}$ denote transaction cost in a neutral network and in a discriminatory network, respectively.*

$$\Delta T \equiv T - \widetilde{T} = \frac{t}{4} - \left(\int_0^{\widetilde{x}} txdx + \int_{\widetilde{x}}^1 t(1 - x)dx\right) = -(\widetilde{x} - \frac{1}{2})^2 t \leq 0 \ \text{for} \ \forall t \ \text{and} \ 1/2 \leq \widetilde{x} \leq 1.$$

∎

Finally, as far as the total delay cost is concerned, we find the following invariance result.

Lemma 6 *The total expected waiting costs are the same in both neutral and discriminatory regimes.*

Proof. We know that the expected waiting cost for each end user in a neutral network is given by $w = \frac{1}{\mu - \lambda}$. With the total number of end users normalized to 1, it also represents the total expected waiting costs denoted by W, i.e., $w = W = \frac{1}{\mu - \lambda}$. The total expected waiting costs in a discriminatory network $\widetilde{W}$ is given by the weighted average costs of

$w_1 = \frac{1}{\mu - \lambda_1}$ and $w_2 = \frac{\mu}{\mu - \lambda} w_1 = \frac{\mu}{\mu - \lambda} \frac{1}{\mu - \lambda_1}$, with weights given by $\widetilde{x}$ and $(1 - \widetilde{x})$, respectively. We also know that $\lambda_1 = \widetilde{x}\lambda$. Thus, we have

$$
\begin{aligned}
\widetilde{W} &= \widetilde{x}w_1 + (1 - \widetilde{x})w_2 \\
&= \frac{\lambda_1}{\lambda} \frac{1}{\mu - \lambda_1} + \left(1 - \frac{\lambda_1}{\lambda}\right) \frac{\mu}{\mu - \lambda} \frac{1}{\mu - \lambda_1} \\
&= \frac{\lambda(\mu - \lambda_1)}{\lambda(\mu - \lambda_1)(\mu - \lambda)} = \frac{1}{\mu - \lambda} = W.
\end{aligned}
$$

∎

As a result, the overall waiting costs are irrelevant in the *static* welfare comparison. Note that this conclusion, however, depends crucially on the assumption that competing contents have the same latency costs. If the latency costs differ across content, the overall waiting costs differ across the regimes.[28]

Considering all three channels through which net neutrality can have an influence upon short-run total welfare, we can conclude that *static* welfare implications of net neutrality regulations depend on the trade-off between transportation cost saving and inefficient production. More specifically, if the production cost asymmetry is quite small, then the production cost effect becomes negligible so that a neutral network would give a higher static social surplus. In contrast, if the production cost difference is significant compared to the transportation cost parameter t, a discriminatory network would be preferred from the social surplus viewpoint. The following proposition summarizes this implication of the net neutrality regulation on social welfare.

Proposition 2 *The comparison of social welfare in the short run with and without net neutrality regulation crucially depends on the relative magnitudes of the production cost asymmetry and the transportation cost parameter. For a sufficiently small asymmetry in production cost, the social welfare is higher under net neutrality, precisely, iff $(c_2 - c_1) < \bar{t}$ where $\bar{t} \equiv \left(\widetilde{x} - \frac{1}{2}\right) \frac{t}{\lambda}$. Otherwise, the discriminatory network yields a higher social surplus.*

The proposition implies that if the two CPs are symmetric in their service cost, the short-run social welfare is higher under net neutrality regulation.

---

[28] In particular, the overall waiting costs would be reduced in a discriminatory regime if the content with higher latency costs is given priority and delivered first. See section 6 for more discussion on this.

Example 1 *Consider the case such as $\mu = 4$, $\lambda = 2$, $t = 1$ so that $\widetilde{x} \approx 0.69$. Assuming $\theta = 0$, for simplicity, the reduction in production cost in the discriminatory network is approximately given by $0.38(c_2 - c_1)$, while the increment in transportation cost is approximately 0.036. Hence, the social welfare under net neutrality is higher than that under discriminatory regime if and only if $0.38(c_2 - c_1) < 0.036$; obviously, for the symmetric production cost of $c_1 = c_2$, net neutrality regulation increases the social welfare.*

4.2 Long Run Analysis with Investment Incentives

Now we extend the analysis to incorporate dynamic considerations such as the broadband operator's incentive to expand capacity in infrastructure. ISPs such as Verizon, Comcast, and AT&T oppose network neutrality regulation claiming that such regulation would discourage their investment incentives in broadband networks. The intuition behind their claims is simple: they face an obvious free-rider problem, unless content providers who support bandwidth-intensive multimedia Internet traffic pay a premium. Here we examine the validity of this claim.

As previously mentioned, we address this issue by investigating the ISP's marginal change in its profit with respect to the capacity parameter μ for the two networks having different governing rules for congestion. Denote $\Phi(\mu)$ to be the cost associated with the capacity level of μ with $\Phi' \geq 0$ and $\Phi'' \geq 0$. Then, the ISP's choice of optimal investment will be determined at the point where the marginal benefit and the marginal cost with respect to μ are equal to each other, i.e., $d\pi_m/d\mu = \Phi'(\mu)$ in the neutral network and $d\widetilde{\pi}_m/d\mu = \Phi'(\mu)$ in the discriminatory network. Note that the marginal benefits of capacity expansion can be written as follows by using the results above:

$$\frac{d\pi_m}{d\mu} = \frac{da}{d\mu} = \frac{1}{(\mu - \lambda)^2} \tag{17}$$

and

$$\frac{d\widetilde{\pi}_m}{d\mu} = \frac{d\widetilde{a}}{d\mu} + \frac{d\widetilde{f}}{d\mu} = \left[\frac{1}{(\mu - \widetilde{x}\lambda)^2} \left(1 - \lambda\frac{d\widetilde{x}}{d\mu} \right) - t\frac{d\widetilde{x}}{d\mu} \right] + 2\left[r - \theta c_1 - (1 - \theta)c_2 \right] \lambda\frac{d\widetilde{x}}{d\mu}. \tag{18}$$

In order to study the condition under which the ISP has a stronger incentive to invest in

the discriminatory network, let us examine the difference between (17) and (18):

$$
\begin{aligned}
\frac{d\widetilde{\pi}_m}{d\mu} - \frac{d\pi_m}{d\mu} &= \left(\frac{d\widetilde{a}}{d\mu} - \frac{da}{d\mu} \right) + \frac{d\widetilde{f}}{d\mu} \\[2mm]
&= \underbrace{\left[\frac{1}{(\mu - \widetilde{x}\lambda)^2} \left(1 - \lambda\frac{d\widetilde{x}}{d\mu} \right) - t\frac{d\widetilde{x}}{d\mu} - \frac{1}{(\mu - \lambda)^2} \right]}_{\substack{\text{changes in the effect of capacity expansion} \\ \text{on end user access fee with discrimination}}} + \underbrace{2\left[r - \theta c_1 - (1 - \theta)c_2 \right] \lambda \frac{d\widetilde{x}}{d\mu}}_{\substack{\text{the effect of capacity expansion} \\ \text{on the sale price of priority right}}}
\end{aligned}
\tag{19}
$$

As can be seen from equation (19), there are two effects in evaluating the relative incentives to invest in capacity across the two regimes.

First, capacity expansion affects the network access fee the ISP can charge end users, which is the willingness to pay by the marginal end users. This *network access fee effect* is represented by the expressions in the square bracket in equation (19). More specifically, in the network with net neutrality, the location of the marginal end user does not change and remains fixed at the midpoint with a change in capacity. However, capacity expansion speeds up the delivery of content uniformly, which enables the ISP to charge more for access. This effect is captured by the last term in the square bracket. In the discriminatory network, capacity expansion affects the delivery speed of content *asymmetrically* across content providers, and thus also changes the location of the marginal consumer type who is indifferent between the two content providers. Such effect of capacity expansion in the discriminatory network is captured by the first two terms in the square bracket. In general, we cannot tell unambiguously the relative size of this network access fee effect under a neutrality regime and under a discriminatory regime: the sign of the square bracketed term in (19) is ambiguous.

Second, capacity expansion also affects the sale price of the priority right under the discriminatory regime. This *rent extraction effect,* represented by the last term in equation (19), weakens the ISP's incentive to invest in capacity under a discriminatory network because the relative merit from first priority and thus its value is relatively small for a higher capacity level. In other words, since the congestion problem becomes less severe for higher capacity levels, the ISP's rent from the allocation of priority classes also decreases, which in turn leads to a weaker investment incentive under a discriminatory regime.

Consequently, the ISP's investment incentive hinges upon the relative magnitudes of

these two potentially opposing effects. It is *a priori* ambiguous whether the ISP has greater incentive to invest in capacity in a neutral network or a discriminatory one. Contrary to the ISPs' claim that net neutrality regulations would have a chilling effect on their incentive to invest, we cannot dismiss the possibility of the opposite. This could happen if capacity expansion alleviates the need to acquire the priority right and hence adversely affects the ability to extract rent from content providers.

Proposition 3 *The ISP's relative incentive to invest in capacity in a discriminatory network vis-a-vis a neutral network depends on two effects: the rent extraction effect and the network access fee effect. The overall effect is ambiguous. In particular, if the rent extraction effect is sufficiently negative, the ISP may invest more on network infrastructure in a neutral network compared to in a discriminatory one.*

Example 2 *Let us consider the same parameter vaules as in Example 1 such that $\mu = 4$, $\lambda = 2$, $t = 1$ and $\theta = 0$. Moreover, we set $r = 1.5$, $c_1 = 0$ and $c_2 = 1$ under which the ISP prefers the discriminatory network. In this case, with some algebra, we derive $\frac{d\tilde{x}}{d\mu} \approx -0.197$ and $\frac{d\tilde{\pi}_m}{d\mu} - \frac{d\pi_m}{d\mu} \approx -0.244$, which numerically demonstrates the possibility that the ISP has weaker investment incentives in discriminatory network.*

One interesting implication of the analysis is that degrading the non-priority packet may be necessary to extract rent more effectively and thus restore incentives to invest in the discriminatory regime. So far, to our best understanding, the opponents of net neutrality have claimed that they have no incentive for degradation even under the discriminatory network.[29] Nevertheless, we must be cautious in interpreting the above proposition. Our result does not necessarily validate the claims from proponents of net neutrality regulation. It just identifies a condition under which the ISP's claim that a discriminatory network is necessary for investment incentives may not be valid.

5   Net Neutrality and CPs' Investment Incentives

So far, our analysis has dealt only with investment incentives of ISPs. As pointed out in von Hippel (2005), proponents of net neutrality regulation maintain that so-called killer

---

[29]For incentives to degrade the quality of a subset of products, see Denecker and McAfee, (1996) and Hahn (2006).

applications have been developed at the 'edges' of the network by users, not by the 'core' of network operators. Thus, another important element in the net neutrality debate is investment incentives for content providers.

5.1 The Hold-up Problem and CPs' Investment Incentives

A typical concern about the so-called hold-up problem is that part of the return from one party's relationship-specific investments is *ex post* expropriable by his trading partner. Such concerns arise when we consider the content service providers' investments: the monopolistic ISP could ex post expropriate any investments made by content providers. The *ex post* optimal policy for ISP to discriminate may not be optimal from an *ex ante* viewpoint. Thus, an interesting question to ask is if the ISP would have the incentive to commit to net neutrality in order to maintain the content providers' incentives to invest.[30]

In order to examine the effect of the discriminatory network on the content providers' R&D incentives, let us assume that a lower marginal cost is achieved at the expense of a higher investment cost. An irreversible investment in cost-reducing R&D is characterized by a twice differentiable function $\Psi(\Delta_i)$ with $\Psi' > 0$, $\Psi'' > 0$, where Δ_i denotes the magnitude of the cost reduction from investing, i.e., $\Delta_i = \bar{c}_i - c_i$. We can think of $\bar{c}_i$ as the current best technology that is freely available to content provider i, and c_i as the post-investment cost level for $i = 1, 2$.

In a neutral network, each content provider's marginal cost reduction increases its profit by $\lambda/2$, which is readily seen from (6). This is because there is no demand effect of cost-reducing investment in the neutral network. Thus, each content provider's optimal investment in cost-reducing R&D is determined by the marginal benefit-cost comparison,

$$\Psi'(\Delta_i^*) = \frac{\lambda}{2} \quad \text{for } i = 1, 2. \tag{20}$$

Similarly, in a discriminatory network each content provider chooses its optimal investment

---

[30]DeGraba (1990) presents a model to study how price discrimination in a market for a variable input affects downstream producers' long-run choices of a production technology. He shows that a monopoly supplier of a variable input will charge the low-cost downstream producer a higher price than the high-cost producer under price discrimination, and thus the downstream producers will end up choosing technology with a higher marginal cost with price discrimination than under uniform pricing, which results in a lower welfare in the long run under discriminatory pricing. using similar reasoning, the literature on the most favored nations (MFN) clause in international trade also suggests that discriminatory or preferential tariffs rather than uniform tariffs would have a more adverse effect on investment incentives of foreign producers (Choi, 1995).

at the point where the marginal revenue from cost-reduction is equalized to the marginal cost. Since the low-cost content provider earns the profit of $\widetilde{\pi}_1^* = (r - c_1)\widetilde{x}\lambda - f$ where f was defined in (12) and the high-cost content provider is not affected by the ISP's rent extraction, content providers' optimal investments are determined by

$$\Psi'(\widetilde{\Delta}_1^*) = (\widetilde{x} - \theta(2\widetilde{x} - 1))\lambda \quad \text{and} \quad \Psi'(\widetilde{\Delta}_2^*) = (1 - \widetilde{x})\lambda. \tag{21}$$

By the comparison of optimal investments under a neutral network with those under a discriminatory one, we derive the following results.

Proposition 4 *The low-cost content provider will choose a technology with a higher marginal cost under the discriminatory network than it will under the neutral network, i.e., $\widetilde{\Delta}_1^* < \Delta_1^*$ if and only if the ISP's expropriation is high enough to the extent of $\theta > 1/2$. Otherwise $(0 \leq \theta \leq 1/2)$, we have $\widetilde{\Delta}_1^* \geq \Delta_1^*$. The high-cost content provider always chooses a technology with a higher marginal cost under the discriminatory network, that is, $\widetilde{\Delta}_2^* < \Delta_2^*$.*

As expected, the optimal investment level of the low-cost content provider is inversely related to the ISP's ability to extract rent from using the fast lane. Suppose that the right to the premium service is traded through the first price bid auction, i.e., $\theta = 0$. Then, the low-cost CP's profit is constrained only by the high cost CP's willingness to pay for the priority service. Since the low-cost CP's cost reduction applies to a larger market coverage in a discriminatory network relative to in neutral network, the low-cost CP will have a stronger investment incentive in a discriminatory regime. Therefore, the low cost CP chooses a technology with a lower marginal cost under a discriminatory regime than under a neutrality regime. Such merit, however, gradually decreases as θ increases. Eventually, for a sufficiently large rent extraction (for $\theta > 1/2$), the low-cost content provider's investment incentive becomes weaker under the discriminatory regime due to rent extraction from the ISP.

On the other hand, the high-cost content provider will always choose a technology with a higher marginal cost under a discriminatory regime for any $\theta \in [0, 1]$. This is because the high-cost content provider always has a smaller market share in the discriminatory network than that in the neutral network. Therefore, the ISP *may* have the incentive to commit to net neutrality to maintain the content providers' innovation incentives.

5.2 Optimal Rent Extraction: Short-Run vs. Long-Run Effect

Discussion in the previous subsection naturally leads us to study the optimal degree of rent extraction in bargaining from the ISP's perspective. Consider a hypothetical situation in which the ISP can choose the parameter θ. Then, we find that there exist interesting intertemporal trade-offs. First, the ISP prefers a larger rent extraction (higher θ) in the short run because of a higher surplus from trading the priority. Had we considered this short run *direct* effect only, the most desirable situation for the ISP is total rent extraction, i.e., $\theta = 1$ with $\frac{\partial f}{\partial \theta} \geq 0$.

From the long-run perspective, however, such total extraction may not be the best option. This is because an increase in its rent extraction can generate the adverse dynamic effect of lowering the low-cost content provider's investment incentive for a higher θ, which in turn can decrease the ISP's long-run revenue from trading the priority.

Therefore, the ISP's optimal level of rent extraction will be determined by these intertemporal trade-offs. To put it mathematically, the overall effect of θ on the ISP's long-run profit is evaluated as

$$
\frac{d\widetilde{\pi}_m^*}{d\theta} \;=\; \underbrace{\frac{\partial \widetilde{\pi}_m^*}{\partial \theta}}_{(+)} + \underbrace{\frac{\partial \widetilde{\pi}_m^*}{\partial \widetilde{\Delta}_1^*}}_{(+)} \underbrace{\frac{\partial \widetilde{\Delta}_1^*}{\partial \theta}}_{(-)}, \tag{22}
$$

where the first term captures the direct rent extraction effect and the second term represents the indirect effect through CP's investment incentives. Needless to say, the ISP will choose θ by $\frac{d\widetilde{\pi}_m^*}{d\theta} = 0$. For an explicit solution, if we consider a quadratic function $\Psi(\Delta_i) = \Delta_i^2/2k$, where k is a cost efficiency parameter in the investment, then the optimal level of θ, denoted by $\tilde{\theta}$, is derived in the following proposition.

Proposition 5 *The ISP's long-run profit is maximized at $\tilde{\theta} = \frac{c_2 - c_1}{(2\tilde{x}-1)k\lambda}$. The ISP does not prefer full rent extraction, if $(0 \leq)\tilde{\theta} = \frac{c_2 - c_1}{(2\tilde{x}-1)k\lambda} < 1$.*

Corollary 1 $\frac{\partial \tilde{\theta}}{\partial k} < 0$ *and* $\frac{\partial \tilde{\theta}}{\partial (c_2 - c_1)} > 0$.

As the content provider's cost-reduction is more efficient (or as parameter k increases), the adverse effect of the ISP's rent extraction on the low-cost content provider's innovation incentive gets large, with all other things being equal. Thus, the ISP's preferred level of

26

rent extraction becomes relatively small. In addition, if the cost differential between the two content providers increases, the ISP will have a stronger incentive to extract more rent from content providers due to the short-run direct effect, *ceteris paribus*.

Example 3 *Consider the case in Example 1, that is, $\mu = 4$, $\lambda = 2$, $t = 1$ so that $\widetilde{x} \approx 0.69$. In addition, let us assume that $c_2 = 2$, $c_1 = 1$, and $k = 5$. Then, the ISP who desires to maximize its profit, with content providers' innovation incentives taken into account, will prefer to have $\tilde{\theta} = \frac{1}{(2*0.69-1)*5} \approx 0.526$.*

In reality, however, the ISP may not have the *ex ante* ability to commit to its preferred level of rent extraction. In such a case, the ISP may prefer to have net neutrality regulations as a commitment device not to extract any rent from CP's investment. The following example shows such a possibility.

Example 4 *Consider the same case as in Example 1, that is, $\mu = 4$, $\lambda = 2$, $t = 1$ so that $\widetilde{x} \approx 0.690$ and $\frac{d\widetilde{x}}{d\mu} \approx -0.197$. In addition, let us assume that $r = 3$, $\bar{c}_1 = 1$, $\bar{c}_2 = 2$ and $\Psi(\Delta_i) = \frac{1}{2}\Delta_i^2$. In a neutral network, the content providers' optimal investments in cost reduction are derived as $\Delta_1^* = \Delta_2^* = 1$, which means that the ex post service costs of content providers are given by $c_1 = 0$ and $c_2 = 1$. By contrast, in a discriminatory network they are derived as $\widetilde{c}_1 = \bar{c}_1 - \widetilde{\Delta}_1^* = 1 - (1.38 - 0.76\theta) = 0.76\theta - 0.38$ and $\widetilde{c}_2 = \bar{c}_2 - \widetilde{\Delta}_2^* = 2 - 0.62 = 1.38$. If we calculate the profit of the ISP across different regulation regimes taking into account this wedge in content providers' investment incentives and the resulting cost levels, the ISP's profit without commitment to network neutrality is given by $\widetilde{\pi}_m^* = v + 1.337\,6\theta - 0.577\,6\theta^2 + 0.159\,52$, but by $\pi_m^* = v - 1$ under net neutrality. Because $\widetilde{\pi}_m^* > \pi_m^*$ for $\forall \theta \in (0,1)$, this example shows that the ISP is able to earn a higher profit under net neutrality regulation than without it.*

6 Discussion and Extensions

6.1 Heterogeneity in Delay Costs across Content

In the basic model, we assumed that the waiting costs due to congestion are identical across content. However, content and applications differ in their sensitivity with respect to delay in delivery. In general, data applications such as email can be relatively insensitive towards

moderate delivery delays from the users' viewpoint. In contrast, streaming video/audio or VoIP applications can be very sensitive to delay, leading to jittery delivery of content. With such heterogeneity concerning delay costs, one may argue that network neutrality treating all packets equally regardless of content is not an efficient way to utilize the network in the presence of capacity constraints. It also has been claimed by opponents of net neutrality regulation that the imposition of net neutrality requirements may impede the development of time-sensitive applications such as remote medical supervision.

To investigate these issues, the model needs to be modified to allow the possibility of different latency costs across applications. More specifically, let us assume τ to be the waiting cost for the low-cost content that would be provided through the fast lane, while that for the high-cost content service is still normalized to one for consistency with the analysis thus far. Because we are particularly interested in the case where the content with higher latency costs is given priority and delivered first, we focus our attention to the case of $\tau \geq 1$.

The marginal consumer who is indifferent between the two content services under the neutrality regime, denoted by x^{**}, is given by

$$x^{**} = \frac{1}{2} + \frac{1 - \tau}{2t(\mu - \lambda)} \leq x^* = 1/2, \tag{23}$$

which means that under net neutrality the demand for the content with higher latency costs decreases compared to the case of identical latency costs. In contrast, under a discriminatory regime the location of the marginal consumer will be given by

$$\widetilde{\widetilde{x}} = \frac{1}{2} - \frac{\tau(\mu - \lambda) - \mu}{2t(\mu - \lambda)(\mu - \widetilde{\widetilde{x}}\lambda)}.$$

By comparing $\widetilde{\widetilde{x}}$ and x^{**}, we find that the low-cost content provider always faces a higher demand for its content service with the first priority relative to in a neutral network, i.e., $\widetilde{\widetilde{x}} > x^{**}$ for any $\tau \geq 1$. The proof of this finding can be readily earned from the facts that the difference between $\widetilde{\widetilde{x}}$ and x^{**}, $\widetilde{\widetilde{x}} - x^{**}$, increases in τ and that $\tau_1 < 1$, where τ_1 is characterized by $\widetilde{\widetilde{x}}(\tau) = x^{**}(\tau)$. Therefore, the qualitative results derived with identical latency costs are quite robust to the relaxation of this assumption *except* with respect to the comparison of social welfare in the short run with and without net neutrality.

Now that there is asymmetry in latency costs across content services, Lemma 4 cannot hold any more. In fact, it becomes possible to have lower total waiting costs under a discriminatory regime relative to those under the neutrality regime if the asymmetry parameter τ is sufficiently high. This is because the effect of the first-priority on the saving of latency costs occurs more favorably toward the content provider who uses the faster lane, while such asymmetrical force disappears with identical latency costs.

Proposition 6 *If $\tau \geq \tau_2 \equiv \frac{\mu(\lambda - \lambda_1) - (1 - x^{**})\lambda(\mu - \lambda_1)}{\lambda(\mu - \lambda_1)x^{**} - \lambda_1(\mu - \lambda)}$ where $\lambda_1 = \widetilde{\widetilde{x}}\lambda$, the total waiting costs are lower under a discriminatory regime than those under a neutral regime.*

As a result, the short-run welfare comparison may move toward favoring the introduction of two-tiered services in the presence of heterogeneity in delay costs across content. Moreover, we find the condition under which one may argue that network neutrality, which treats all packets equally regardless of content, is not an efficient way to utilize the network in the presence of capacity constraints.

6.2 Possibility of Quality Degradation

As we pointed out early on, the ISP may want to degrade the quality of non-priority packets (deliberately slow down the delivery speed of content) for the purpose of extracting rent more effectively and restoring incentives to invest in the discriminatory regime. In this spirit let us consider quality degradation for the basic service by allowing the ISP to be able to choose a waiting time higher than w_2 in (3) for non-priority packets.

We find that the ISP can have incentive to do quality degradation in a discriminatory network, but not in a neutral network. This is because in a neutral network the ISP's quality degradation only decreases the network access fee without yielding a higher rent extraction. Secondly, as is obvious from the ISP's profit in the discriminatory network in (13), the low-cost content provider will have a larger market share with such quality degradation than without it. The enlarged asymmetry in the demands for content can make the ISP earn more from the trade of the first priority to the low-cost content provider, but reduce the ISP's revenue from the network access fee. As long as the former effect outweighs the latter, the possibility of quality degradation would make discriminatory network more profitable for ISPs.

29

Once again, a question of interest is how the possibility of quality degradation affects the investment incentives of the ISP. With the possibility of quality degradation, the ISP need not to be concerned anymore about the rent extraction effect that adversely affects the ISP's incentive to invest in capacity expansion. In other words, the ISP is now free of the problem that the relative quality difference between the two CPs decreases as capacity expands. Thus, the possibility of quality degradation can increase ISPs' incentives to expand capacity.

6.3 Integration/Strategic Alliance of ISPs and CPs

Another important issue in the debate on net neutrality is the impact of integration of ISPs and content providers on market competition and innovation incentives. One concern expressed by net neutrality proponents is the possibility that the integrated ISPs may confer unfair advantage to its own content over content provided by competitors. Consider, for instance, a recent merger of AT&T with SBC that has a partnership with Yahoo. The question is whether AT&T would have an incentive to give its partner Yahoo site preferential treatment over competing sites such as Google in the absence of net neutrality regulations. To address this question, we need to analyze whether the ISP may have incentives to offer the first-priority to the affiliated content provider over the non-affiliated one.

In our simple model, it turns out that under net neutrality vertical integration has no impact on allocation of resources either in the short-run or in the long-run. Therefore, there is no antitrust concern about vertical merger between the ISP and CP: if there is a vertical merger, it is driven by efficiency reasons. Even without net neutrality, it can be shown that the allocation of the first-priority is the same across different vertical structures in that the low cost CP always receives the first priority. Therefore, the concern that the ISP may give its own sister division preferential treatment over competing sites is unfounded at least in the short-run. However, vertical integration in a discriminatory regime can have impacts on capacity investment of the ISP. To see this, let us consider a vertical merger between the ISP and the low cost CP and denote the merged firm's profit as $\widetilde{\Pi} = \widetilde{a} + (r - c_1)\widetilde{x}\lambda$, where $\widetilde{a} = v - \frac{1}{\mu - \widetilde{x}\lambda} - t\widetilde{x}$. Then, the merged firm's investment incentives can be expressed as

$$\frac{d\widetilde{\Pi}}{d\mu} = \frac{d\widetilde{a}}{d\mu} + [r - c_1]\lambda\frac{d\widetilde{x}}{d\mu}. \tag{24}$$

Notice that the merged firm's investment incentives do not depend on θ, because the sale of the first priority is internal to the organization.[31]   By comparing (24) and (18), the comparison of investment incentives with vertical integration and without vertical integration depends on the relative magnitude of $2\left[r - \theta c_1 - (1-\theta)c_2\right]$ and $[r - c_1]$. Noting that $\frac{d\widetilde{x}}{d\mu} < 0$, the ISP's investment incentives with vertical integration are higher than those under no vertical integration if θ is sufficiently high and close to 1. The reason is that with vertical integration the ISP does not need to deliberately limit its capacity in an effort to command a higher sale price for the first priority. However, if θ is sufficiently small, the result can be reversed. More specifically, if $r - c_2 < (c_2 - c_1)(1 - 2\theta)$, an independent ISP has higher incentives to invest than a vertically merged one. Note that this condition is identical to the one that ensures that CP1 benefits from a discriminatory regime. This condition holds when the independent ISP's ability to extract rent from the sale of the first priority is limited and thus ISP does not fully internalize the negative impact of capacity investment on the relative value of first priority. Once integrated, it fully internalizes its impact on CP1's profit and thus limits its investment to confer advantage to its own CP division.

7   Concluding Remarks

This paper provides an economic analysis of the net neutrality regulation. In particular, our analysis focuses on the effects of net neutrality regulation on the investment incentives of ISPs and CPs as well as on social welfare. To address these questions, we use a simple model based on the queuing theory to capture the congestion in the network. We have shown that the ISP's incentives to invest in a multi-tiered network vis-a-vis in a nondiscriminatory network under net neutrality regulation depends on a potential trade-off between the two-sides of the market: the network access fee from the end users and the revenue from content providers through the potential trade of the first-priority in delivery.   We also compare the CPs' incentives to invest in cost reduction/quality enhancement as well as social welfare across different regulatory regimes. We find that the relationship between the net neutrality regulation and investment incentives is subtle. Even though we cannot draw general unambiguous conclusions, we identified key effects that are expected to play

---

[31]If the merger took place between the ISP and the high cost CP, the incentive to invest will depend on θ.

important roles in the assessment of net neutrality regulations.

We conclude by mentioning some limitations of our simple model and discussing potential avenues for future research. First, we note that the model in the previous sections made many simplifying assumptions with regard to pricing strategies of several players. For instance, we assumed that the ISP does not charge content providers under net neutrality regulations and charges only the content provider who purchases premium services in a discriminatory network. In general, the ISP can charge content providers under net neutrality with the restriction that they are charged the same price without any priority in service. We also assumed away the ability of content providers to charge end users directly. Consideration of these possibilities considerably complicates the analysis. In this regard, the burgeoning literature on two-sided markets may be useful in further analyzing these issues.[32] In the framework of two-sided markets, ISPs will play the role of platforms that provide a link between content providers and end users. Caillaud and Jullien (2003), for instance, show that the equilibrium in two-sided markets depends crucially on the pricing scheme used. Thus, it would be important to analyze the implications of allowing a more sophisticated pricing scheme in this model. In particular, it would be an important extension to allow competition between content providers when micropayments between content providers and consumers are possible.

Second, one may consider introducing diversity in the types of investments that can be made by content providers. More specifically, we can imagine two types of investments: firm-specific investments, whose effects are limited to the investing content providers, and investments that have spillover effects. For the first type of investment, we can think of investments that enhance the value of content or reduce the cost of content provision. For the second type, we can consider an investment in compression technology, which not only reduces the delivery speed of the investor's content, but relieves congestion in the network that helps delivery speed of other content providers. The net neutrality regulations may have a differential effect across different types of investments and impact the choice of investments.

Finally, our basic framework assumes that the ISP market is characterized by monopoly power. This is a reasonable approximation in many geographical markets. However, it is not the only market condition prevailing. One important extension of the model would be

---

[32]See Armstrong (2006) and Rochet and Tirole (2006) for details.

to introduce competition in the ISP market and analyze how the effects of net regulation can play out. Most concerns expressed by net neutrality proponents are rooted in the monopoly power and concentration in the ISP market. One important policy question would be whether the presence of competition in the ISP market can mitigate any problems associated with discrimination and make net neutrality regulation irrelevant.

References

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

[2] Balachandran, K.R., "Purchasing Priorities in Queues," *Management Science*, vol. 18(5), 1972, pp. 319-326.

[3] *Business Week Online*, "Rewired and Ready for Combat," November 7, 2005, available at http://www.businessweek.com/magazine/content/05_45/b3958089.htm.

[4] Caillaud, Bernard and Bruno Jullien, "Chicken and Egg: Competition among Intermediation Service Providers," *Rand Journal of Economics*, Summer 2003, pp. 309-328.

[5] Cerf, Vinton G., "Prepared Statement in the U.S. Senate Committee on Commerce, Science, and Transportation Hearing on 'Net Neutrality'," February 7, 2006.

[6] Cheng, Hsing Kenneth, Subhajyoti Bandyopadhyay, and Hong Guo, "The Debate on Net Neutrality: A Policy Perspective," unpublished manuscript, 2006.

[7] Chirico, Van de Haar and Larouche, "Network Neutrality in the EU," TILEC, discussion paper, 2007.

[8] Choi, Jay Pil, "Optimal Tariffs and the Choice of Technology: Discriminatory Tariffs vs. the 'Most Favored Nation' Clause," *Journal of International Economics*, January 1995, pp. 143-160.

[9] Choi, Jay Pil, "Tying in Two-Sided Markets with Multi-Homing," NET Institute Working Paper #06-04, available at http://www.netinst.org/Choi2006.pdf.

[10] Deneckere, Raymond J. and R. Preston McAfee, "Damaged Goods," *Journal of Economics & Management Strategy*, vol. 5(2), 1996, pp. 149-174.

[11] Economides, Nicholas, "Net Neutrality, Non-Discrimination and Digital Distribution of Content through the Internet," NET Institute Working Paper #07-03, 2007, available at http://www.stern.nyu.edu/networks/Economides_Net_Neutrality.pdf.

[12] Economides, Nicholas and Joacim Tåg, "Net Neutrality on the Internet: A Two-sided Market Analysis" NET Institute Working Paper No. 07-14, NYU Law and Economics Research Paper

[13] Edelson, Noel M. and David K. Hilderbrand, "Congestion Tolls for Poisson Queuing Processes," *Econometrica*, vol. 43 (1), January 1975, pp. 81-92.

[14] Gross, Donald and Carl M. Harris, Fundamentals of Queueing Theory, 3rd Ed. Wiley., 1998.

[15] Hahn, Jong-Hee, "Damaged Durable Goods," *RAND Journal of Economics*, vol. 37(1), Spring 2006, pp. 121-133, .

[16] Hermalin, Benjamin E. and Michael L. Katz, "The Economics of Product-Line Restrictions with an Application to the Net Neutrality Debate," AEI-Brookings Joint Center for Regulatory Studies Working Paper 07-02, February 2007.

[17] Kocsis, Viktória and Paul W.J. de Bijl, "Network Neutrality and the Nature of Competition between Network Operators," *International Economics and Economic Policy*, vol. 4, 2007, special issue on the Digital Economy and Regulatory Issues, pp. 159-184.

[18] Mendelson, Haim, "Pricing Computer Services: Queueing Effects," *Communications of the ACM*, vol. 28(3), 1985, pp. 312-321.

[19] Mendelson, Haim. and S. Whang, "Optimal Incentive-Compatible Priority Pricing for the M/M/1 Queue," *Operations Research*, vol. 38, 1990, pp. 870-883.

[20] Naor, P., "The Regulation of Queue Size by Levying Tolls," *Econometrica*, vol. 37(1), January 1969, pp. 15-24.

[21] Rochet, Jean-Charles and Jean Tirole, "Two-Sided Markets: A Progress Report," *Rand Journal of Economics*, vol. 37(3), Autumn 2006, pp. 645-667.

[22] Valletti, Tommaso and Carlo Cambini, "Investments and Network Competition," *Rand Journal of Economics*, vol. 36(2), 2005, pp. 446-468.

[23] Van Schewick, Barbara, "Towards an Economic Framework for Network Neutrality Regulation," *Journal on Telecommunications and High Technology Law 5,* 2007, pp.329-392.

[24] Von Hippel, Eric, Democratizing Innovation, 2005, MIT Press.

[25] Wu, Tim, "Network Neutrality, Broadband Discrimination," *Journal on Telecommunications and High Technology Law 2,* 2003, pp.141-178.

[26] Yoo, Christopher S., "Network Neutrality and the Economics of Congestion," *Georgetown Law Journal*, vol. 94, August 2006, pp. 1847-1908.

Appendix

Proof of Proposition 4. Recall that $\Psi'(\Delta_i^*) = \frac{\lambda}{2}$, $\Psi'(\widetilde{\Delta}_1^*) = (\widetilde{x} - \theta(2\widetilde{x} - 1))\lambda$ and $\Psi'' > 0$. Thus, $\Delta_1^* > \widetilde{\Delta}_1^*$ if and only if $\frac{1}{2} > \widetilde{x} - \theta(2\widetilde{x} - 1)$. Because $\frac{\partial}{\partial\theta}(\widetilde{x} - \theta(2\widetilde{x} - 1)) = -(2\widetilde{x} - 1) < 0$, the condition for $\widetilde{x} - \theta(2\widetilde{x} - 1) < \frac{1}{2}$ is equal to the condition for $\widetilde{x} - \theta(2\widetilde{x} - 1) - \frac{1}{2} = (2\widetilde{x} - 1)(\frac{1}{2} - \theta) < 0$. Hence, $\theta > \frac{1}{2}$ is the necessary and sufficient condition for $\Delta_i^* > \widetilde{\Delta}_1^*$. Similarly, the comparison between $\Psi'(\Delta_2^*) = \frac{\lambda}{2}$ and $\Psi'(\widetilde{\Delta}_2^*) = (1 - \widetilde{x})\lambda$ yields the result of $\widetilde{\Delta}_2^* < \Delta_2^*$. ∎

Proof of Proposition 5. The marginal revenue from an increase in θ, the first term in (22), is given by $\frac{\partial\widetilde{\pi}_m^*}{\partial\theta} = (c_2 - c_1)(2\widetilde{x} - 1)\lambda$ from (13). Note that the marginal cost-reduction of the low-cost content provider due to a marginal increase in θ is given by $\frac{\partial\widetilde{\Delta}_1^*}{\partial\theta} = -(2\widetilde{x} - 1)k\lambda$ and the feedback effect of cost-reduction on the ISP's profit is derived as $\frac{\partial\widetilde{\pi}_m^*}{\partial\widetilde{\Delta}_1^*} = \theta(2\widetilde{x}-1)\lambda$. Thus, the boomerang effect is measured by $\frac{\partial\widetilde{\pi}_m^*}{\partial\widetilde{\Delta}_1^*}\frac{\partial\widetilde{\Delta}_1^*}{\partial\theta} = -\theta(2\widetilde{x}-1)^2k\lambda^2$. Thus, the overall marginal effect of the degree of rent extraction on the ISP's payoff is derived as:

$$
\begin{aligned}
\frac{d\widetilde{\pi}_m^*}{d\theta} &= (c_2 - c_1)(2\widetilde{x} - 1)\lambda - \theta(2\widetilde{x} - 1)^2k\lambda^2 \\
&= (2\widetilde{x} - 1)\lambda\left[(c_2 - c_1) - \theta(2\widetilde{x} - 1)k\lambda\right],
\end{aligned}
$$

from which we can see $\frac{d\widetilde{\pi}_m^*}{d\theta} = 0$ at $\tilde{\theta} = \frac{c_2 - c_1}{(2\widetilde{x}-1)k\lambda}$. ∎

Proof of Proposition 6. With the heterogeneity in delay costs, in a discriminatory network we have $w_1 = \frac{\tau}{\mu - \widetilde{\widetilde{x}}\lambda}$ and $w_2 = \frac{\mu}{\mu - \lambda}\frac{1}{\mu - \widetilde{\widetilde{x}}\lambda}$. Thus, the total expected waiting costs are respectively given by

$$
\begin{aligned}
\widetilde{\widetilde{W}} &= \frac{\lambda_1}{\lambda}\frac{\tau}{\mu - \widetilde{\widetilde{x}}\lambda} + \left(1 - \frac{\lambda_1}{\lambda}\right)\frac{\mu}{\mu - \lambda}\frac{1}{\mu - \widetilde{\widetilde{x}}\lambda} \\
&= \frac{\lambda_1\tau(\mu - \lambda) + (\lambda - \lambda_1)\mu}{\lambda(\mu - \lambda_1)(\mu - \lambda)} \quad \text{where } \lambda_1 = \widetilde{\widetilde{x}}\lambda
\end{aligned}
$$

and

$$
W^{**} = x^{**}\frac{\tau}{\mu - \lambda} + (1 - x^{**})\frac{1}{\mu - \lambda} = \frac{1 + x^{**}(\tau - 1)}{\mu - \lambda}.
$$

The condition of $W^{**} \geq \widetilde{\widetilde{W}}$ is equal to that of

$$
\begin{aligned}
\left(1 + x^{**}(\tau - 1)\right) \lambda(\mu - \lambda_1) \quad &\geq \quad \lambda_1 \tau(\mu - \lambda) + (\lambda - \lambda_1)\mu \\
\Longleftrightarrow \quad \tau &\geq \tau_2 \equiv \frac{\mu(\lambda - \lambda_1) - (1 - x^{**})\lambda(\mu - \lambda_1)}{\lambda(\mu - \lambda_1)x^{**} - \lambda_1(\mu - \lambda)}
\end{aligned}
$$

∎