RAND Journal of Economics
Vol. 41, No. 3, Autumn 2010
pp. 446–471

Net neutrality and investment incentives

Jay Pil Choi*

and

Byung-Cheol Kim**

This article 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. Concerning the ISPs' investment incentives, we find that capacity expansion decreases the sale price of the priority right under the discriminatory regime. Thus, 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. A discriminatory regime can also weaken CPs' investment incentives because of CPs' concern that the ISP would expropriate some of the investment benefits.

1. Introduction

■ This article analyzes the effects of net neutrality regulation on investment incentives for various players in the Internet market. Since the Internet's inception, one of the governing principles in its operation has been nondiscrimination 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 nondiscrimination restrictions. In fact, major telephone and cable operators, which together control about 98% of broadband service in the United States (as of December 2005),[1] recently expressed an interest in providing multitier Internet service, charging content providers (CPs) premium prices for preferential access to broadband transmission service. In response, a coalition of content providers merged forces in an effort to maintain the nondiscrimination status

*University of New South Wales, Australia; choijay@gmail.com.
**Georgia Institute of Technology; byung-cheol.kim@econ.gatech.edu.
We are grateful to the coeditor and two anonymous referees for their valuable comments and suggestions, which greatly improved the article. We also thank Paul de Bijl, Nick Economides, and participants in various conferences and seminars for helpful comments, and Jane Choi for her editorial assistance. The first draft of this article was completed during Choi's research visit to the ISCR and the School of Economics and Finance at Victoria University of Wellington, whose hospitality is greatly appreciated. This article received the Public Utility Research Center Prize for the best paper in regulatory economics, which was awarded at the Seventh Annual International Industrial Organization Conference. We gratefully acknowledge financial support from the NET Institute for this research. All errors, if any, are ours.

[1] FCC Form 477 data.

quo. Their intensive lobbying efforts led to a hot debate—known as the net neutrality debate—in Washington, along with initiatives to legislate a mandate to prevent creating a multitier Internet service. Although efforts to enact net neutrality regulations have stalled 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 United States' largest cable TV operator and second largest Internet provider, had interfered with users' access to file-sharing sites such as BitTorrent.[3] This practice is 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." As one person's upload is another's download in file-sharing networks, this type of traffic management can have a series of repercussions in the network of file sharers. As a result, the incident received nationwide attention and stirred uproar from users of file-sharing applications who were adversely affected.

To inform this important policy debate, we analyze economic issues associated with net neutrality regulation. Considering that the Internet is a vital medium of communication, information, and commercial activities, it is of paramount importance to maintain competition and promote innovation in this market. 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.

Reflecting the importance of the Internet as a main driver of economic growth and prosperity in the global economy, one of the main issues surrounding 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 supporting bandwidth-intensive multimedia applications pay a premium for heavy Internet traffic. By 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 nondiscriminatory neutrality principle since its earliest days. To support their claim that net neutrality has been the main driver of the growth and innovative applications of the Internet, 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" (Cerf, 2006). 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 based on the queuing theory because 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.[5] With a monopolistic network operator and two application providers, we provide a formal economic analysis of the effects of net neutrality regulations on investment incentives for ISPs and CPs, and

[2] The Obama administration has expressed support for net neutrality. On October 22, 2009, the FCC released the Net Neutrality NPRM (Notice of Proposed Rulemaking). In particular, Paragraph 106 of the NPRM states that "a broadband Internet access service provider may not charge a content, application, or service provider for enhanced or prioritized access to the subscribers of the broadband Internet access service provider." For detailed explanation and discussion of institutional differences between the European Union and United States concerning net neutrality regulation, see Chirico, Van de Haar, and Larouche (2007).

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

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

[5] For an economic analysis of traffic congestion at the interface between backbones in the commercial Internet, see Cremer, Rey, and Tirole (2000).

their implications for social welfare. There is no universally accepted definition of net neutrality. In principle, net neutrality means that all packets that traverse through the Internet are treated equally on the basis of first come, first served. Any routing practice deviating from this principle is a violation of net neutrality, for example, port blocking, quality degradation, and access tiering. In our article, we focus on prioritization in delivery speed for particular packets as the main mode of nonneutrality. Without net neutrality regulation, it is possible for a content provider to pay for preferential delivery while best efforts are assured for the rest without targeted degrading or purposeful content blocking.[6]

We 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 with the equilibrium in which it cannot discriminate the delivery speed of content. This comparison of two short-run equilibria yields two major findings. First, in a discriminatory network, 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. The ISP's decision of whether or not it will prefer the discriminatory regime to the neutral network depends on a potential tradeoff 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 the discriminatory regime if the asymmetry across content providers is sufficiently large.

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 an 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 magnitudes of these two potentially opposing effects. 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 the 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

[6] The reason for assuming best efforts for basic service is to reflect the current climate in the debate on net neutrality. No ISPs contemplate engaging in such a deliberate degradation of service, perhaps in fear of regulatory backlash. We discuss the possibility of deliberate degradation in Section 7. See Kocsis and de Bijl (2007) for various types of violations of net neutrality.

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 article is organized in the following way. The next section offers a brief literature review of articles addressing net neutrality issues. Section 3 sets up a basic model of network markets to analyze the effects of net neutrality regulation on competition and social welfare. Section 4 analyzes the short-run equilibrium with the capacity fixed and studies the effects of net neutrality on ISPs, CPs, consumer surplus, and social welfare. Sections 5 and 6 analyze the effects of net neutrality regulations on investment incentives of ISPs and CPs, respectively. In Section 7, 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 ISPs and CPs. Section 8 closes this article with concluding remarks, along with suggestions for further possible extensions of our basic analysis. The proofs of lemmas and propositions are relegated to Appendix A.

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.[7]

Hermalin and Katz (2007) examine a situation in which ISPs serve as a platform to connect content providers with end consumers in a framework of two-sided markets. They consider heterogeneous content providers whose products are vertically differentiated to analyze the effects of net neutrality regulation. Without any restrictions, ISPs can potentially offer a continuum of vertically differentiated services, although the ISP is required to provide only one service (a single tier of Internet service) with net neutrality regulation. They compare the single-service level equilibrium with the multiservice level equilibrium and 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 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 effects of such regulation can be ambiguous, but they argue that the effects are often negative. Our research complements the analysis of Hermalin and Katz (2007) in that we consider the congestion effect in the provision of Internet service and, more importantly, investigate investment incentives of ISPs and CPs; both are not addressed in theirs.

In terms of policy questions and basic framework, our research is closest to Cheng, Bandyopadhyay, and Guo (2009), who develop a game-theoretic model of competition between two content providers in a Hotelling framework. They investigate the effects of net neutrality regulation on ISPs' incentives to expand capacity as well as study who gains and who loses as a consequence of the regulation. We build upon their framework and extend their analysis in several directions. However, there are several major differences between the two articles. First, our study goes one step further by analyzing the effects of the regulation on content providers' incentives to provide innovative services. We find that the holdup problem can prevail under a discriminatory regime, and thus *ex ante* the ISP might prefer not to extract the entire rent available from CPs. Second, our model adopts a different pricing scheme in the sale of the priority. They assume that the ISP deals with the two competing CPs nonexclusively and charges both for priority regardless

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

of what the rival CP does. In such a setup, it is possible that both CPs pay the price for priority, but end up being back where they started because no CP would have an advantage over the other. Our model, by contrast, assumes that the ISP deals with CPs exclusively in that only one CP can be given priority. This different consideration about the pricing scheme leads to more general results with new insights. For instance, Cheng et al. find that the ISP's incentive to expand its capacity is unambiguously higher under net neutrality, but such an outcome is just one possibility in our model. In fact, we discover 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 addition, they find that if the principle of net neutrality is abandoned, the ISPs definitely stand to gain from extracting fees for preferential treatment from the content providers. In our model, however, such a relationship is not always derived, because the discrimination may decrease the ISP's revenue from network access fees for consumers.

Economides and Tåg (2007) also provide an economic analysis of net neutrality in a two-sided market framework, but their research focus differs from ours. They are particularly interested in how net neutrality regulation affects pricing schemes on both sides of the market and social welfare in the short run; thus, our study strongly complements theirs.

Finally, Valletti and Cambini (2005) analyze network operators' incentives to invest in networks with different quality levels, as in our article. They show that quality has an impact on all calls initiated by customers (destined both on-net and off-net) and that "tacit collusion" takes place even in a symmetric model with two-part pricing, because firms tend to underinvest in quality. 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 content to end consumers through a broadband network provided by a monopolistic Internet service provider. 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. As Economides (2008) argues, the monopolistic ISP assumption is reasonable in the United States when we consider limited choices of broadband Internet access available to residential consumers as well as significant switching costs in changing ISPs (see Kocsis and de Bijl, 2007 for competition between oligopolistic network operators).

To understand our notion of net neutrality and the issues we address, it is useful to note that the Internet is an interconnected "network of networks." None of the network operators has the capacity to provide complete end-to-end routing between content providers and end consumers. Typically, the network operator who provides hosting services to a content provider would be different from the ISP who provides Internet connection to an end consumer. Thus, when a consumer requests specific content from a content provider, it needs to traverse several different networks. When the content is delivered, even packets comprising the same web page can travel different routes before they are assembled at the client's computer. The transit between networks is governed by a variety of peering agreements between networks. Tier 1 networks constitute the Internet backbone and have direct connections to the Internet. Tier 1 network operators interconnect with each other without purchasing transit or paying settlements. Tier 2 and 3 networks are relatively small players and purchase at least some transit from other networks to reach the Internet. We take these existing peering agreements between networks as given.

Our analysis focuses on the local market in which a local monopolistic ISP provides the "last-mile" connection service to consumers. In particular, we study the effects of the local ISP's discretion to discriminate content once the packets from content providers have arrived at the local ISP's switch box through the existing transit agreement between networks. In other words, we consider the possibility of two-tier service at the local ISP level with the ISP extracting additional payment from content providers as a price for "priority" delivery in the absence of net neutrality regulation.

□ **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 and does not charge content providers for sending information over its network to end users (Hahn and Wallsten, 2006).[8] However, note that this does not mean that CPs receive access to the Internet for free under net neutrality. We envision a typical situation in which the network operators that provide hosting service to CPs are different from the local ISP that provides access to end users. Under net neutrality, CPs pay access fees to the hosting network operator only once at the origin and are not required to pay additionally for "transit," which is already covered in the existing traffic-routing arrangement governed by peering agreements between networks. Thus, the local ISPs that provide the last-mile transit to end consumers are not allowed to demand additional compensation from CPs. By 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 of the two content providers.

We assume that each consumer has demands for one of two CPs. The content request rate of each consumer follows a Poisson process with hazard rate λ, which represents the demand intensity. The network capacity is denoted by μ. The service time taken to deliver the content from CPs to end users is exponentially distributed with its mean of $1/\mu$. Larger network capacity implies a shorter service time. In the short-run analysis, the capacity μ is assumed to be fixed. In the long-run analysis in which investment incentives are investigated, it is endogenously derived.

Consumers, whose mass is normalized to one, are heterogeneous with respect to their preferences toward two content services in the Hotelling manner. Setting CP1 and CP2 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. The transport cost per unit distance, t, can represent the degree of product differentiation. As in Mendelson (1985), we assume that each consumer with content request rate λ derives a gross utility of $u(\lambda) = v$ from either content service, and that v is sufficiently large so that the market is fully covered in both regimes of networks.[9]

As in Cheng et al. (2009), Choi (2010), and Economides and Tåg (2007), we assume that content providers adopt a business model that offers their services without any direct charge to consumers but generates revenues through advertisements. Each content provider i earns a revenue stream r_i from advertisers for each consumer's content request ("click-throughs") it serves. The asymmetry in r_i may reflect differences in CPs' capabilities to match advertisers and consumers. The cost of serving each consumer's request is given by c_i. Content provider i's markup per each consumer's click-through is given by $m_i = r_i - c_i$, where $m_1 \geq m_2 \geq 0$ without loss of generality. Thus, the sources of asymmetry in CPs' click-through margins can be either from the revenue side, the cost side, or a combination of both. The corresponding CP's profit is measured by $m_i \lambda \sigma_i$, where σ_i denotes the market share for content provider i.

The sequence of the players' choices is as follows. In the discriminatory network regime, the ISP can first sell the priority service through a trading process to only one content provider exclusively; in the neutral network, this stage does not apply. Then, the ISP posts a network access fee, a, to end users. Given the allocation of the priority classes and the network access fee, end users choose one of the content providers. As usual, the analysis for this game proceeds by using

[8] 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.

[9] The content request rate λ 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. We treat such a possibility as a second-order effect and ignore it.

backward induction, and the equilibrium concept employed here is that of subgame-perfect Nash equilibrium.

□ **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; Balachandran, 1972; Edelson and Hilderbrand, 1975; Mendelson and Whang, 1990). The reason for this modelling choice is two-fold. First, the root causes of the net neutrality debate are scarce bandwidth and the potential need for rationing due to substantial increases in bandwidth-intensive multimedia applications over the Internet. Second, 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 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'. Furthermore, this microfoundation yields nice properties with which we can analyze without any ad hoc assumptions.

In the neutral network where all packets are treated equally without any priority classes, each consumer has the expected waiting time of

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

where λ denotes the gross content request rate at the network (with the normalization of consumer mass to one) with the network capacity $\mu > \lambda$. The waiting time increases in λ, but decreases in μ. If we normalize the delay cost per unit time to one, then the waiting time in (1) equals each consumer's expected waiting cost. In the basic model, we assume that all content has the same delay cost per unit time; in Section 7, we extend our analysis by considering heterogeneity in delay costs across content and applications.

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 nonpreemptive discriminatory network, 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 the first priority.[10] By contrast, the consumer who requests content without the 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,[11] we can infer that a consumer will face a higher waiting cost by requesting the nonprioritized content instead of the prioritized one, that is,

Property 1. $w_2 > w > w_1$ for $\mu > \lambda$.

Property 1 is easily established by examining the relative ratio of w_2 to w_1, that is, $w_2/w_1 = \mu/(\mu - \lambda) > 1$. As a related property, we can also notice 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

[10] See Gross and Harris (1998) and references therein for more on the queuing theory and for the detailed derivation for waiting costs in different types of networks.

[11] In discriminatory networks, there are two possible priority schemes: preemptive and nonpreemptive. 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 nonpreemptive 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 progress.

addition, we find that the quality difference measured in waiting costs becomes smaller as the network capacity increases, that is,

$$\text{Property 2.} \quad \frac{\partial}{\partial \mu}(w_2 - w_1) < 0.$$

This is because the marginal reduction in waiting time for the fast lane from capacity expansion decreases as the capacity level becomes high.

We adopt the M/M/1 system for our analysis because it is a standard framework to model congestion in computer networks. However, the same qualitative results can be derived with a more general framework as long as the two properties above are satisfied.

4. Net neutrality and short-run analysis

□ **Equilibrium in the neutral network.** In a neutral network, end users choose one of the two content providers who provides higher net surplus, knowing that the waiting cost is given by (1). In the Hotelling model of end users, the marginal consumer x^* who is indifferent between two content providers in the neutral network is defined as[12]

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

where consumers whose preferences are represented by $x < x^*$ choose CP1 and those by $x > x^*$ 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, that is, $x^* = 1/2$. We further consider a scenario in which each consumer's taste parameter x is fixed, which implies that it is only the middle consumer ($x = 1/2$) whose participation constraint is binding.[13]

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 access fee and each content provider's profit as

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

□ **Equilibrium in the discriminatory network.** If the ISP is allowed to charge content providers for the higher priority class, consumers will face different expected waiting times depending on their choice of content service, as derived in (2) and (3). Let us assume that the high-margin (more efficient) content provider, CP1, obtains the first priority, with its content being entitled to be served ahead of CP2's. Later, we demonstrate that this scenario arises as an equilibrium outcome regardless of the trading mechanism. The marginal consumer at $\widetilde{x}$, who is indifferent between the premium service and the basic service, is characterized by the following equality:

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

[12] 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 exploration for the implications of direct payment will be an important extension of this basic model, as explained in Section 8.

[13] Alternatively, we can imagine a situation in which each consumer has a random arrival rate of demands for CP service, and for each demand, their taste parameter x is i.i.d. rather than constant for all time. In such a specification, each consumer would have *ex ante* the same preference and thus the same expected surplus from using the ISP. This alternative specification yields qualitatively the same results as in our specification.

where we use a tilde to denote variables associated with the discriminatory regime.[14] By comparing (4) and (7), we can derive an intuitive result that the content provider with the first priority has a larger market share than the one without it, that is, $\widetilde{x} \geq x^* = 1/2$, due to the difference in waiting costs.

Note that as more consumers switch from CP2 to CP1, the delivery speed of CP1's premium service deteriorates, but the delivery speed of CP2's basic service worsens more.

$$\frac{\partial w_2}{\partial \lambda_1} = \frac{\mu}{\mu - \lambda} \frac{1}{(\mu - \lambda_1)^2} > \frac{1}{(\mu - \lambda_1)^2} = \frac{\partial w_1}{\partial \lambda_1} > 0 \tag{8}$$

The gap in waiting time between the premium and basic services widens as more consumers switch from CP2 to CP1, which in turn makes CP1 more attractive. This positive-feedback process can lead to a corner solution, a situation in which all consumers subscribe to the CP with the first priority. To ensure an interior equilibrium in which the CP with the first priority does not corner the market ($\sigma_i > 0$ for $i = 1, 2$), we make the following assumption.

Assumption 1. $t > \frac{\lambda}{(\mu-\lambda)}2$.

The assumption says that the content of the two CPs are sufficiently differentiated to prevent a corner solution.[15]

The following lemma specifies a sufficient condition under which an interior market-sharing equilibrium is stable and the market share of the content provider with the first priority decreases as the ISP's capacity increases.

Lemma 1. Assuming that $\mu > \frac{3\lambda}{2}$, we have a stable interior equilibrium in a discriminatory network with $\widetilde{x} \in (1/2,\ 1)$. The market share of the CP with content delivery priority decreases as the ISP's capacity increases, that is, $\frac{d\widetilde{x}}{d\mu} < 0$.

The main intuition for this result is that an increased capacity of an ISP makes congestion less important and reduces the relative quality differential (i.e., difference in waiting costs) across the two CPs. In the rest of the article, we assume that $\mu > \frac{3\lambda}{2}$ to focus on the stable equilibrium. Our analysis thus proceeds with $\frac{d\widetilde{x}}{d\mu} < 0$.

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

$$\max_{\widetilde{a}} \widetilde{\pi}_m = \widetilde{a} + f \qquad \text{s.t.} \qquad v - \frac{1}{\mu - \widetilde{x}\lambda} - t\widetilde{x} - \widetilde{a} \geq 0, \tag{9}$$

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. As each content provider knows that its market share will be $\widetilde{x}$ if it acquires the priority right but $(1 - \widetilde{x})$ if the other CP acquires the priority, its maximum willingness to pay for the priority is given by $m_i(2\widetilde{x} - 1)\lambda$. For instance, if the priority right is sold through a first-price ascending auction, CP1 will receive the priority with $f = m_2(2\widetilde{x} - 1)\lambda$, which is CP2's maximum willingness to pay for the right.[16] Alternatively, we can 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, it makes another offer to CP2. In such a scenario, the ISP can extract all surplus from CP1 by charging $f = m_1(2\widetilde{x} - 1)\lambda$.

We adopt a framework that can encompass both scenarios above and the full range between them to represent different surplus divisions between the ISP and the CP that acquires the priority. To encompass the full range of bargaining protocols including the above two extremes,

[14] The explicit formula for $\widetilde{x}$ is given by $\widetilde{x} = \frac{1}{4\lambda}(2\mu + \lambda - \sqrt{(2\mu - \lambda)^2 - \frac{8\lambda^2}{t(\mu-\lambda)}})$.

[15] See the proof of Lemma 1 in Appendix A to see that this condition ensures an interior solution.

[16] Economides (2008) 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.

let $\theta (0 \leq \theta \leq 1)$ denote the ISP's bargaining power that measures the proportion of rent extraction from CP1.[17] Then, the price of the first priority will be given as

$$\begin{aligned} f &= \theta m_1(2\widetilde{x} - 1)\lambda + (1 - \theta)m_2(2\widetilde{x} - 1)\lambda \\ &= (m_2 + \theta \Delta_m)(2\widetilde{x} - 1)\lambda, \end{aligned} \tag{10}$$

where $\Delta_m = m_1 - m_2 (\geq 0)$ and $\theta \in [0, 1]$. As expected, the more bargaining power the ISP has, the higher the priority price will be, that is, $\frac{\partial f}{\partial \theta} = \Delta_m(2\widetilde{x} - 1)\lambda \geq 0$.

Therefore, we can express the ISP's profit in the discriminatory network as

$$\widetilde{\pi}_m^* = \left(v - \frac{1}{\mu - \widetilde{x}\lambda} - t\widetilde{x}\right) + (m_2 + \theta \Delta_m)(2\widetilde{x} - 1)\lambda. \tag{11}$$

When the ISP assigns the priority to CP1 at the price in (10), each content provider's profit will be respectively given by

$$\widetilde{\pi}_1^* = m_1\widetilde{x}\lambda - (m_2 + \theta \Delta_m)(2\widetilde{x} - 1)\lambda; \qquad \widetilde{\pi}_2^* = m_2(1 - \widetilde{x})\lambda. \tag{12}$$

If the two margins are equal, then the extent of ISP bargaining power is irrelevant. In the discriminatory regime both CPs—the one with and the one without priority—make the same profit.

□ **The effects of net neutrality on ISPs' profits.** Here we analyze the effects of net neutrality regulation on the ISP's profits by comparing (6) and (11). We find the following potential tradeoff: 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 one content provider (f).

Lemma 2. The network access fee in the discriminatory network is lower than that in the neutral network, that is, $\widetilde{a} < a^*$.

Lemma 2 reflects Property 1 ($w_2 > w > w_1$) and the associated result $\widetilde{x} \geq x^*$. In the discriminatory network, the network access fee the ISP can charge to end users is reduced. In the absence of net neutrality regulation, the ISP will choose to introduce a two-tiered service if its revenue from the priority trade outweighs the loss from the reduced network access fee.[18]

Proposition 1. The ISP's incentives to introduce the discriminatory network can be summarized as follows.

(i) If $m_2 > \Lambda$, $\widetilde{\pi}_m^* > \pi_m^*$ for all $\theta \in [0, 1]$.
(ii) If $m_1 < \Lambda$, $\widetilde{\pi}_m^* < \pi_m^*$ for all $\theta \in [0, 1]$.
(iii) If $m_2 \leq \Lambda \leq m_1$, there exists a critical level of $\theta^* \in [0, 1]$ such that $\widetilde{\pi}_m^* > \pi_m^*$ iff $\theta > \theta^*$, where $\Lambda = \frac{t}{2\lambda} + \frac{1}{(2\widetilde{x}-1)\lambda}\left(\frac{1}{\mu - \widetilde{x}\lambda} - \frac{1}{\mu - \lambda}\right)$.

Proposition 1 identifies the beneficiaries and losers of net neutrality regulation. Part (i) states that the ISP's profit is always higher in the discriminatory network if the advertising revenue margins (m_i) are sufficiently high for both CPs. By contrast, if the advertising revenue margins are relatively low for both CPs, the ISP would prefer a neutral network. Finally, if the advertising margin is high for one CP but low for the other, the relative merits of the discriminatory network vis-à-vis the neutral network depend on the ISP's bargaining power against the CPs.

The parameter m_i represents the importance of market share for each CP. As m_i's are increased, CPs compete more aggressively to obtain the first priority in the discriminatory network. The ISP thus receives a higher price for the premium service, which can outweigh any potential loss in access fees from end users. This also implies that if both m_is are sufficiently low or only

[17] We do not pin down detailed microfoundations for the bargaining process, because such an issue is not the focus of our article.

[18] Even if we consider the possibility that the ISP charges CPs in the net neutrality regime and the losing CP in the discriminatory regime, our qualitative results do not change.

m_2 is sufficiently low but the ISP's bargaining power is low, the ISP will endogenously choose the equal treatment of both content providers even if net neutrality is *not* required.

Cheng et al. (2009) also consider the effects of a discriminatory network on the ISP's profits, and it is useful to compare our results to theirs. They adopt a different pricing scheme for the sale of the priority. More specifically, Cheng et al. assume that the ISP deals with the two competing CPs in a nonexclusive way and charges a price for priority regardless of what the rival CP does. As a result, it is possible that both CPs pay but with the end result that both CPs are on equal footing without any CP having an advantage over the other. Our model, by contrast, assumes that the ISP deals with CPs exclusively. However, a direct comparison of their results to ours is somewhat difficult. The reason is that our model considers a more general trading mechanism that can encompass a variety of bargaining environments, whereas Cheng et al. consider a mechanism in which the ISP makes a nonexclusive but take-it-or-leave-it offer. Thus, in our notation, they consider only the case where $\theta = 1$. In Cheng et al., the ISP is always better off with a nondiscriminatory regime because it has the option of charging a small amount for "priority" to induce both CPs to pay, without affecting the access fee that the ISP can charge to consumers. This outcome, however, depends crucially on the assumption that the priority right is sold only once. Imagine a more realistic case where content providers can quickly respond by purchasing their own priority right when the delivery of their content is disadvantaged, which is possible when the priority right is sold in a nonexclusive way. In such a case, a more reasonable outcome would be a "delayed-purchase" equilibrium in which no content provider purchases the priority right, as in Fudenberg and Tirole's (1987) model of technology adoption. Then, offering priority in an exclusive way would be the optimal policy for the ISP. In Appendix B, we compare the relative profitability of exclusive and nonexclusive sale of priority rights for the ISP, assuming that the ISP can commit to a one-time sale of the priority right if selling nonexclusively.

□ **The effects of net neutrality on CPs' profits and consumer welfare.** We now turn our attention to the effects of a discriminatory network on CPs and consumers.

Proposition 2. The effects of a discriminatory network on CPs and consumers are as follows.

(i) CP1: $\widetilde{\pi}_1^* > \pi_1^*$ if $\frac{m_1}{m_2} > \frac{2(1-\theta)}{1-2\theta}$ and $\theta \in [0, 1/2)$; otherwise, $\widetilde{\pi}_1^* < \pi_1^*$.
(ii) CP2: $\widetilde{\pi}_2^* \leq \pi_2^* \forall m_i, \theta, \lambda$.
(iii) Users: aggregate consumer welfare increases.

By comparing (6) and (12), we find that the high-margin CP has a higher profit in the discriminatory network if $\theta \in [0, 1/2)$ and $\frac{m_1}{m_2} > \frac{2(1-\theta)}{1-2\theta}$. The best-case scenario for the high-margin CP with priority is to capture the whole market and double the profit it would receive under net neutrality. If $\theta > 1/2$ and thus more than half of its profit is extracted, it cannot be better off under the discriminatory regime. When $\theta < 1/2$, the high-margin content provider who obtains the first priority can have a higher payoff in the discriminatory regime if the margin ratio between the two content providers is sufficiently large. Note that the threshold $\frac{2(1-\theta)}{1-2\theta}$ is an increasing function of $\theta \in [0, 1/2)$, which implies that the margin ratio needs to be greater for the high-margin CP to earn more profit in the discriminatory network as the ISP extracts more rent from it. Because $\frac{2(1-\theta)}{1-2\theta}$ reaches its minimum value of 2 for $\theta = 0$, we can conclude that the necessary condition for the high-margin CP to prefer the discriminatory network is $\theta < 1/2$ and $\frac{m_1}{m_2} > 2$.

By contrast, the low-margin CP is always worse off from the introduction of priority classes. These results show the possibility that both content providers may engage in a Prisoners' Dilemma type of game to receive the first priority but end up with lower payoffs, whereas the ISP prefers the discriminatory network. Both CPs can also be worse off in Cheng et al. (2009). Such an outcome, however, takes place only when both CPs purchase the priority right, which is excluded in our model. If only one CP pays for the priority delivery, the paying CP is indifferent between the two regimes in their model.

Finally, the result in (iii) states that the end users as a group enjoy a higher surplus in the discriminatory network. This clean result stems from our Hotelling specification where an individual consumer's surplus increases linearly with the distance of her location from the marginal consumer who is indifferent between the two CPs. Hence, the aggregate consumer surplus in the discriminatory network is greater than that with net neutrality in which the two CPs share the market evenly, thereby minimizing the total consumer surplus.

□ **The effects of net neutrality on 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 benefits/costs we need to compare to analyze the effects of two-tiered pricing on social welfare: (i) total margins for CPs, (ii) total transportation costs, and (iii) total delay costs. In our welfare analysis, we assume that CPs' revenue r_i from advertisers reflects the social benefits of advertising.[19] The following lemma examines the effects of these factors on the short-run social welfare one by one.

Lemma 3. (i) The total margins for CPs are larger under the discriminatory regime relative to the neutral one. (ii) The total transportation costs are higher under the discriminatory regime relative to the neutral one. (iii) The total expected waiting costs are the same in both regimes.

The series of the results in Lemma 3 is very intuitive. First, the discriminatory regime allows the high-margin content provider to expand its market share through speedier delivery of its content. As a result, efficiency in terms of margin maximization favors the discriminatory network. Second, recalling that the total transportation costs are minimized when the marginal consumer is located at the midpoint, the two-tiered pricing with $\widetilde{x} > 1/2$ is inefficient in terms of transportation cost minimization. Finally, as far as the total delay cost is concerned, we find the invariance result as shown in Appendix A. The simple reason is that the change in regimes only affects the order of services but not the total amount of services in the network. Given a fixed network capacity, the amount required to serve all users should be the same, which implies identical average waiting times across the two regimes. As a result, the overall waiting costs are irrelevant in the *static* welfare comparison. This conclusion, however, depends crucially on the assumption that competing contents have the same waiting costs. If the waiting costs differ across content, the overall waiting costs differ across the regimes (see Section 7 for more discussion).

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 tradeoff between transportation cost saving and inefficient production. If the margin difference is significantly large relative to the degree of product differentiation, the discriminatory network would be preferred from the viewpoint of social welfare.

Proposition 3. The comparison of social welfare in the short run with and without net neutrality regulation crucially depends on the relative magnitudes of the margin asymmetry across CPs (Δ_m) and the transportation cost parameter. If the margin difference is significantly large relative to the degree of product differentiation, the social welfare is higher in the discriminatory network, precisely, iff $\Delta_m > \bar{t} \equiv (\widetilde{x} - \frac{1}{2})\frac{t}{\lambda}$.

The proposition implies that if the two CPs are symmetric in their margins ($\Delta_m = 0$), the short-run social welfare is higher under net neutrality regulation.

5. Long-run analysis with investment incentives

■ The net neutrality debate centers on future investment and innovations.[20] In particular, one of the main issues in the debate is how the broadband operator's incentive to expand capacity in

[19] See Anderson and Coate (2005) for a microfoundation of such informative advertising.

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

infrastructure would be affected by allowing preferential transmission of content. ISPs such as Verizon, Comcast, and AT&T oppose network neutrality regulation and claim 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 Internet traffic pay a premium.

We examine the validity of this claim by investigating the marginal change in the ISP's profit with respect to the capacity parameter μ for the two network regimes. 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, that is, $d\pi_m/d\mu = \Phi'(\mu)$ 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:

$$\frac{d\pi_m}{d\mu} = \frac{da}{d\mu} = \frac{1}{(\mu - \lambda)^2} \tag{13}$$

and

$$\frac{d\widetilde{\pi}_m}{d\mu} = \frac{d\widetilde{a}}{d\mu} + \frac{df}{d\mu} = \left[\frac{1}{(\mu - \widetilde{x}\lambda)^2}\left(1 - \lambda\frac{d\widetilde{x}}{d\mu}\right) - t\frac{d\widetilde{x}}{d\mu}\right] + 2(m_2 + \theta\Delta_m)\lambda\frac{d\widetilde{x}}{d\mu}. \tag{14}$$

Therefore,

$$\frac{d\widetilde{\pi}_m}{d\mu} - \frac{d\pi_m}{d\mu} = \left(\frac{d\widetilde{a}}{d\mu} - \frac{da}{d\mu}\right) + \frac{df}{d\mu}$$

$$= \underbrace{\left[\frac{1}{(\mu - \widetilde{x}\lambda)^2}\left(1 - \lambda\frac{d\widetilde{x}}{d\mu}\right) - t\frac{d\widetilde{x}}{d\mu} - \frac{1}{(\mu - \lambda)^2}\right]}_{\substack{\text{changes in the effect of capacity expansion}\\ \text{on end user access fee due to discrimination (?)}}} + \underbrace{2(m_2 + \theta\Delta_m)\lambda\frac{d\widetilde{x}}{d\mu}}_{\substack{\text{the effect of capacity expansion}\\ \text{on the sale price of priority right (}-\text{).}}} \tag{15}$$

As can be seen from equation (15), there are two effects to consider when 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 brackets in equation (15). 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. 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 brackets, $\frac{1}{(\mu-\lambda)}^2$. By contrast, 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 an effect of capacity expansion in the discriminatory network is captured by the first two terms in the square brackets. The first term, $\frac{1}{(\mu-\widetilde{x}\lambda)^2}(1 - \lambda\frac{d\widetilde{x}}{d\mu})$, measures the effect of capacity expansion on the consumer's network access fee through the improved delivery speed of content. This effect can be further decomposed into two forces. The first part, $\frac{1}{(\mu-\widetilde{x}\lambda)^2}$, measures the increase in the marginal consumer's willingness to pay for network access when he subscribes to the CP with priority, with the demand configuration between the two CPs fixed. Note that the benefit from a larger capacity when he subscribes to the CP with priority, $\frac{1}{(\mu-\widetilde{x}\lambda)^2}$, is less than $\frac{1}{(\mu-\lambda)^2}$, which is the benefit from a capacity expansion when no CP has priority. The reason is that when a CP has priority and its content is already delivered quickly, the beneficial effect of capacity expansion on delivery speed is relatively small. However, there is a secondary effect from capacity expansion that goes in the opposite direction. Whereas the demand configuration is fixed under net neutrality, capacity expansion under the discriminatory regime favors the CP without priority. (Recall Lemma 1, which shows $\frac{d\widetilde{x}}{d\mu} < 0$.) Thus, capacity expansion induces demand reduction for the CP with priority and thus further diminishes potential congestion for the content with the first priority. The increment in the marginal user's willingness to pay due to this demand effect is captured by the second part, $-\lambda \frac{1}{(\mu-\tilde{x}\lambda)^2} \frac{d\tilde{x}}{d\mu} (> 0)$. In addition, the capacity expansion decreases the transportation cost of the marginal consumer who subscribes to the CP with priority in the discriminatory regime (once again, recall $\frac{d\tilde{x}}{d\mu} < 0$). Such savings in transportation costs will also increase the marginal consumer's willingness to pay for network access, which is captured by the second term of the square brackets, $-t\frac{d\tilde{x}}{d\mu} (> 0)$. Because $\frac{1}{(\mu-\tilde{x}\lambda)^2} < \frac{1}{(\mu-\lambda)^2}$, the sign of the square bracketed term in (15) is indeterminate, so we cannot tell unambiguously the relative size of this network access fee effect under the neutral regime and under the discriminatory regime.

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 (15), 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, because 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.

In general, 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. Nonetheless, the analysis in this article unveils what forces can make ISPs' innovation incentives strong or weak in each regime. Although our model allows for the possibility that the ISP can have stronger investment incentives without net neutrality regulation, we cannot dismiss the possibility of the opposite. Contrary to the ISPs' claim that net neutrality regulations would have a chilling effect on their incentive to invest, we find that net neutrality regulations could rather boost the incentive for ISP capacity expansion because it alleviates the need to acquire the priority right and hence adversely affects the ability to extract rent from content providers.

Proposition 4. The ISP's relative incentive to invest in capacity in a discriminatory network vis-à-vis a neutral network depends on two effects: the rent extraction effect and the network access fee effect. The overall effect is ambiguous.

To understand the conditions under which the ISP may have more incentives to invest under net neutrality, we can rearrange the terms in (15) as

$$\frac{d\tilde{\pi}_m}{d\mu} - \frac{d\pi_m}{d\mu} = \underbrace{\left[\frac{1}{(\mu-\tilde{x}\lambda)^2} - \frac{1}{(\mu-\lambda)^2}\right]}_{\substack{\text{direct effects of capacity expansion} \\ \text{with market shares fixed}\,(-)}} + \underbrace{\underbrace{\left[2(m_2+\theta\Delta_m) - \frac{1}{(\mu-\tilde{x}\lambda)^2} - \frac{t}{\lambda}\right]}_{\equiv\Sigma}\lambda\underbrace{\frac{d\tilde{x}}{d\mu}}_{(-)}}_{\substack{\text{indirect effects of capacity expansion} \\ \text{through changes in market shares.}}} \tag{16}$$

The terms in the first square brackets represent the direct effects of capacity expansion on the difference in the ISP's profits between the two regimes with the CPs' market shares *fixed*, which is always negative. The remaining terms represent the effects of capacity expansion through the *induced changes* in the CPs' market shares. Thus, a sufficient condition for the ISP to have higher incentives to invest in capacity is that $\Sigma \geq 0$. Thus, as CPs' margins are higher and the ISP's bargaining power becomes stronger, it becomes more likely that the ISP will have more incentives to invest in a neutral network. The reason is that the rent extraction motives are stronger under such a situation. Another scenario under which the ISP may have higher incentives to invest in capacity occurs when $\frac{d\tilde{x}}{d\mu}$ is close to zero. This scenario takes place if the product differentiation parameter t is sufficiently high so that the indirect effects through changes in market shares are negligible. In this scenario, the direct effect dominates and the ISP will have higher incentives to invest under net neutrality.

6. 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 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.

A typical concern about the so-called holdup 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 expropriate any investments made by content providers. The *ex post* optimal policy for the ISP to discriminate may not be optimal from an *ex ante* investment incentive viewpoint. Thus, an interesting question to ask is whether regulation is required as a mechanism to bind the ISP to net neutrality in order to maintain the content providers' incentives to invest.[21]

In order to examine the effect of the discriminatory network on the content providers' R&D incentives, let us assume that a higher margin is achieved with a higher investment cost.[22] An irreversible investment in margin-improving 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 margin enhancement from investing, that is, $\Delta_i = m_i - \bar{m}_i$. We can think of $\bar{m}_i$ as the current margin with the best technology that is freely available to content provider i, and m_i as the postinvestment margin level for $i = 1, 2$.

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

$$\Psi'(\Delta_i^*) = \frac{\lambda}{2} \quad \text{for } i = 1, 2. \tag{17}$$

Similarly, in a discriminatory network, where CPs decide their investment levels followed by the competition for priority, each content provider chooses its optimal investment at the point where the marginal revenue from margin improvement is equalized to the marginal cost. One complication in a discriminatory network, however, is that the marginal benefit from investment for a CP depends on whether or not it receives the priority, which is endogenously determined by the comparison of CPs' *ex post* margins after investment. If the initial margin difference between CP1 and CP2 is sufficiently large, the unique equilibrium entails that CP1 always has a higher *ex post* margin and thus receives the priority. In this case, the high-margin content provider earns a profit of $\tilde{\pi}_1^* = m_1\tilde{x}\lambda - f$ where f was defined in (10), and the low-margin content provider is not affected by the ISP's rent extraction. Thus, content providers' optimal investments are given by

[21] 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).

[22] The investment can be either revenue enhancing or cost reducing.

$\widetilde{\Delta}_1^* = \bar{\Delta}$ and $\widetilde{\Delta}_2^* = \underline{\Delta}$, where $\bar{\Delta} \geq \underline{\Delta}$ (equality holding when $\theta = 1$) with $\bar{\Delta}$ and $\underline{\Delta}$ being defined by

$$\Psi'(\bar{\Delta}) = (\widetilde{x} - \theta(2\widetilde{x} - 1))\lambda \quad \text{and} \quad \Psi'(\underline{\Delta}) = (1 - \widetilde{x})\lambda. \tag{18}$$

If the initial margin difference is small, we cannot rule out the possibility that the *ex post* margin ranking can be reversed in equilibrium. However, the qualitative results do not change, as demonstrated below.

Lemma 4. Under the discriminatory network, one CP invests at the level of $\bar{\Delta}$ whereas the other CP invests at the level of $\underline{\Delta}$ in any pure strategy equilibrium.

Thus, in any pure strategy equilibrium, one CP invests $\bar{\Delta}$ whereas the other invests $\underline{\Delta}$.[23] In other words, the identity of the CP that invests more and receives the priority may change, but the overall equilibrium investment levels do not change in a discriminatory network. Thus, we will focus on the pure strategy equilibrium where the high-margin firm invests more and retains the priority *ex post*. By comparing optimal investments under a neutral network with those under a discriminatory one, we derive the following results.

Proposition 5. The high-margin content provider will choose a technology with a lower margin under the discriminatory network than it will under the neutral network, that is, $\widetilde{\Delta}_1^* < \Delta_1^*$, if and only if the ISP's expropriation is high enough to the extent of $\theta > 1/2$. Otherwise (if $0 \leq \theta \leq 1/2$), we have $\widetilde{\Delta}_1^* \geq \Delta_1^*$. The low-margin content provider always chooses a technology with a lower margin under the discriminatory network, that is, $\widetilde{\Delta}_2^* < \Delta_2^*$.

As expected, the optimal investment level for the high-margin content provider is inversely related to the ISP's ability to extract rent from use of the fast lane. Suppose that the right to the premium service is traded through the first-price bid auction, that is, $\theta = 0$. Then, the high-margin CP's profit is constrained only by the low-margin CP's willingness to pay for the priority service. Because the high-margin CP's margin improvement applies to larger market coverage in the discriminatory network relative to in the neutral network, the high-margin CP will have a stronger investment incentive in a discriminatory regime. Therefore, the high-margin CP chooses a technology with a higher margin under a discriminatory regime than under a neutral regime. Such merit, however, gradually decreases as θ increases. Eventually, for a sufficiently large rent extraction (for $\theta > 1/2$), the high-margin content provider's investment incentive becomes weaker under the discriminatory regime due to rent extraction from the ISP.[24]

On the other hand, the low-margin content provider will always choose a technology with a lower margin under a discriminatory regime for any $\theta \in [0, 1]$. This is because the low-margin content provider always has a smaller market share in the discriminatory network than in the neutral network. This implies that the ISP may have a dynamic inconsistency problem when CPs have opportunities to invest either in revenue-enhancing or cost-reducing R&D that improves their margins. For instance, we can imagine a situation in which $\bar{m}_1 < \Lambda$, and thus the ISP *ex ante* prefers a neutral network (see Proposition 1), but once the innovation takes place the ISP prefers to switch to a discriminatory network, that is, $\bar{m}_2 + \Delta_2^* > \Lambda$. This situation would apply to an emerging technology for which the initial margins are low but the potential for technology improvement is huge. Of course, if such *ex post* opportunistic behavior by the ISP is anticipated by

[23] There can also be a mixed-strategy equilibrium at the investment stage where the two firms randomize between $\bar{\Delta}$ and $\underline{\Delta}$.

[24] Our analysis has assumed that the CP's investment raises its margin. Alternatively, we can consider a scenario in which CP investment raises the consumers' gross utility from content service v, which we assumed to be constant and symmetric across CPs. This alternative specification complicates the analysis because we need to account for the channel through which investment in quality helps the CP gain market share. Nonetheless, we can derive the same qualitative results as in Proposition 5.

the CPs, their investment will be adversely affected for fear of rent extraction.[25] In such a scenario, net neutrality regulation may be needed to restore CPs' investment incentives. This result accords well with the fact that content providers in general are the proponents of net neutrality regulation and the crux of their main argument concerns their innovation incentives.

The discussion above also has some implications for the optimal degree of rent extraction in the discriminatory network from the ISP's long-run perspective. The ISP has the following intertemporal tradeoffs. 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, that is, $\theta = 1$ 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 high-margin 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

$$\frac{d\widetilde{\pi}_m^*}{d\theta} = \underset{(+)}{\frac{\partial \widetilde{\pi}_m^*}{\partial \theta}} + \underset{(+)}{\frac{\partial \widetilde{\pi}_m^*}{\partial \widetilde{\Delta}_1^*}} \underset{(-)}{\frac{\partial \widetilde{\Delta}_1^*}{\partial \theta}}, \tag{19}$$

where the first term captures the direct rent extraction effect and the second term represents the indirect effect through CPs' 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 $\widetilde{\theta}$, is derived in the following proposition.

Proposition 6. The ISP's long-run profit is maximized at $\widetilde{\theta} = \frac{\Delta_m}{(2\widetilde{x}-1)k\lambda}$. The ISP does not prefer full rent extraction if $\widetilde{\theta} = \frac{\Delta_m}{(2\widetilde{x}-1)k\lambda} \in [0, 1)$. The ISP's optimal level of rent extraction is decreasing in the CPs' R&D efficiency, but increasing in the margin differential, that is, $\frac{\partial \widetilde{\theta}}{\partial k} < 0$ and $\frac{\partial \widetilde{\theta}}{\partial \Delta_m} > 0$.

As the content provider's R&D process is more efficient (or as parameter k increases), the adverse effect of the ISP's rent extraction on the high-margin content provider's innovation incentive becomes greater, with all other things being equal. Thus, the ISP's preferred level of 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*.

7. Discussion and extensions

□ **Heterogeneity in delay costs.** 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. By contrast, streaming video/audio or voice over Internet protocol (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.

[25] One question that can be asked is why the ISP and the CPs cannot write a contract to solve this holdup problem. We can easily imagine a situation in which the magnitude of margin improvements is observable by the ISP, but not verifiable in court. Then it cannot be included in the contract and the holdup problem cannot be mitigated through an appropriate contract.

To investigate these issues, the model needs to be modified to allow the possibility of different waiting costs across applications. More specifically, let us assume τ to be the waiting cost per unit time for the high-margin content service that would be provided through the fast lane, whereas that for the low-margin content service is still set to one per unit time for consistency with the analysis thus far. Because we are particularly interested in the case where the content with higher waiting costs is given priority and delivered first, we focus our attention on the case of $\tau \geq 1$.[26]

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

$$v - \frac{1}{\mu - \lambda}\tau - tx^{**} - a = v - \frac{1}{\mu - \lambda} - t(1 - x^{**}) - a. \tag{20}$$

Thus, we have

$$x^{**} = \frac{1}{2} - \frac{\tau - 1}{2t(\mu - \lambda)} \leq x^{*} = 1/2, \tag{21}$$

which means that under net neutrality the demand for the content with higher waiting costs decreases compared to the case of identical waiting costs. Similarly, under the discriminatory regime the location of the marginal consumer will be given by[27]

$$\approx{x} = \frac{1}{2} - \frac{\tau(\mu - \lambda) - \mu}{2t(\mu - \lambda)(\mu - \approx{x}\lambda)}. \tag{22}$$

By comparing $\approx{x}$ and x^{**}, we find that the high-margin content provider always faces a higher demand for its content service with the first priority relative to in the neutral network, that is, $\approx{x} > x^{**}$ for any $\tau \geq 1$. This finding can be readily derived by the fact that the difference between $\approx{x}$ and x^{**}, $\approx{x} - x^{**}$, increases in τ and that $(\approx{x} - x^{**})|_{\tau=1} = \tilde{x} - x^{*} > 0$. Therefore, the qualitative results derived with identical waiting 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 the asymmetry in waiting costs across content services, Lemma 3 (iii) does not hold anymore. Assigning priority to content with high waiting costs is certainly beneficial in reducing total waiting costs. However, we cannot conclude that we have lower total waiting costs under the discriminatory regime relative to those under the neutral regime. That would be true if the market shares between the two CPs were the same across the regimes. However, giving priority to CP1's content in a discriminatory regime leads to a higher market share of content with higher waiting costs ($\approx{x} > x^{**}$). This indirect market demand effect can offset the direct effect of allocating priority service to content with higher waiting costs. In fact, our simulation exercises indicate that in most cases the induced demand effect dominates the direct effect, and thus the total waiting costs in fact increase with priority service.

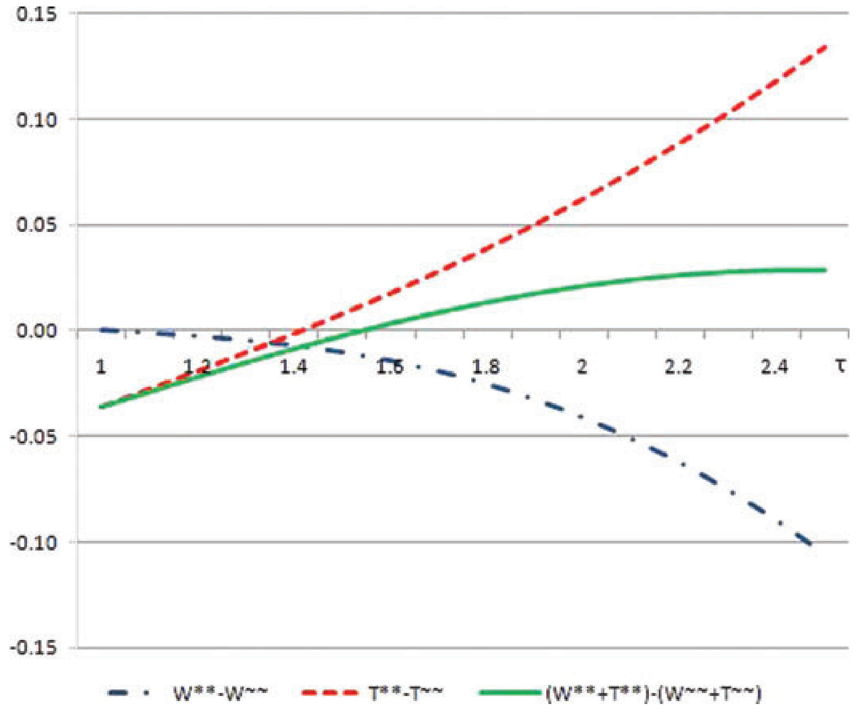
However, this does not imply that priority service in the discriminatory network reduces social welfare. As the asymmetry in waiting costs (τ) increases, the total transportation costs favor the discriminatory network compared to the neutral network. To see this, note that as τ is increased from 1, the total transportation costs start to increase in the neutral regime (x^{**} departs further from 1/2) whereas they decrease in a discriminatory regime ($\approx{x}$ converges to 1/2 as τ reaches $\frac{\mu}{\mu-\lambda}$). Beyond the point of $\tau = \frac{\mu}{\mu-\lambda}$, the total transportation costs increase in both regimes, but still the transportation costs increase at a faster rate in the neutral network. 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. Thus, one may argue that

[26] The high-margin CP is willing to pay more for priority, but this does not necessarily mean that the CP with the greater delay cost will purchase the priority. If $\tau < 1$, it is possible that the content with the smaller delay cost is delivered with first priority.

[27] The explicit formula for $\approx{x}$ is given by $\approx{x} = \frac{1}{4\lambda}((\lambda + 2\mu) - \sqrt{(2\mu - \lambda)^2 + \frac{8\lambda(\tau(\mu-\lambda)-\mu)}{t(\mu-\lambda)}})$.

FIGURE 1

THE EFFECTS OF WAITING-COST ASYMMETRY ON SOCIAL WELFARE



network neutrality, which treats all packets equally regardless of content, is not an efficient way to utilize the network in the presence of heterogeneity in delay costs.

To illustrate our discussion above, we provide a simulation result. By Proposition 3, we know that social welfare is higher in the neutral network with the assumption that both CPs have the same margin ($m_1 = m_2$) for their click-throughs. Our simulation result shows that the social welfare can be higher in a discriminatory regime if τ is sufficiently large even with $m_1 = m_2$, which confirms our intuition.[28]

Figure 1, based on simulation results in Table 1, illustrates that a discriminatory network may perform better (i.e., incur lower total costs) relative to a neutral network for a sufficiently large τ due to the effects discussed above.

□ **Possibility of quality degradation.** One interesting implication of our analysis is that degrading the nonpriority packet may be necessary to extract rent more effectively and thus restore the ISP's incentives to invest in the discriminatory network. So far, to our best understanding, the opponents of net neutrality have claimed that they have no incentive for degradation even under the discriminatory regime.[29] Let us discuss this inquiry in our framework of whether the ISP has incentives to degrade the quality of nonpriority packets—deliberately slow down their delivery speed—for the purpose of extracting rent more effectively and restoring incentives to invest in the discriminatory network. Quality degradation for basic service can be easily modelled by allowing the ISP to choose a waiting time higher than w_2 in (3) for nonpriority packets.

[28] If content with higher waiting costs provides a higher margin, this will further favor a discriminatory network.

[29] This may be due to the ISP's fear of public backlash that would provide impetus for net neutrality regulation once such an intention is revealed. For incentives to degrade the quality of a subset of products, see Deneckere and McAfee (1996) and Hahn (2006). See also Cremer, Rey, and Tirole (2000) for incentives to selectively degrade the quality of interconnection between Internet backbone providers.

TABLE 1 A Numerical Simulation the Effects of Waiting-Cost Asymmetry

	Demands			Waiting Costs		Transport Costs		Difference in Total Costs
τ	$\tilde{\tilde{X}}$	X^{**}	$\tilde{\tilde{X}} - X^{**}$	$\tilde{\tilde{W}}$	W^{**}	$\tilde{\tilde{T}}$	T^{**}	$(W^{**} + T^{**}) - (\tilde{\tilde{W}} + \tilde{\tilde{T}})$
1.0	0.691	0.500	0.191	0.500	0.500	0.286	0.250	−0.036
1.2	0.648	0.450	0.198	0.548	0.545	0.272	0.253	−0.022
1.4	0.608	0.400	0.208	0.587	0.580	0.262	0.260	−0.009
1.6	0.570	0.350	0.220	0.620	0.605	0.255	0.273	0.003
1.8	0.534	0.300	0.234	0.646	0.620	0.251	0.290	0.013
2.0	0.500	0.250	0.250	0.667	0.625	0.250	0.313	0.021
2.2	0.467	0.200	0.267	0.683	0.620	0.251	0.340	0.026
2.4	0.436	0.150	0.286	0.695	0.605	0.254	0.373	0.028

Note: All parameters satisfy stable market-sharing equilibrium conditions ($\mu = 4,\ \lambda = 2, t = 1$).

We find that the ISP can have incentive to do quality degradation in the discriminatory network, but not in the neutral network. This is because in the neutral network the ISP's quality degradation only decreases the network access fee without yielding a higher rent extraction. In addition, as is seen from (11), the high-margin 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, but reduces 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 the discriminatory network more profitable for ISPs.

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 be concerned anymore about the rent extraction effect that adversely affects its investment incentives to capacity expansion. Because the ISP is now free of the problem that the relative quality difference between the two CPs decreases as capacity expands, the possibility of quality degradation would increase ISPs' incentives to expand capacity.

□ **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 advantages to their own content over content provided by competitors. Consider, for instance, a recent merger of AT&T with SBC which 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 nonaffiliated 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 mergers 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 high-margin 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, a vertical integration in the discriminatory regime can have impacts on the ISP's capacity investment. To see this, let us consider a vertical merger between the ISP and the high-margin CP, and denote the merged firm's profit as $\widetilde{\Pi} = \widetilde{a} + m_1\widetilde{x}\lambda$, where $\widetilde{a} = v - \frac{1}{\mu - \widetilde{x}\lambda} - t\widetilde{x}$. Then, the merged firm's investment incentives can be expressed as

$$\frac{d\widetilde{\Pi}}{d\mu} = \frac{d\widetilde{a}}{d\mu} + m_1\lambda\frac{d\widetilde{x}}{d\mu}. \tag{23}$$

Notice that the merged firm's investment incentives in capacity (μ) do not depend on θ, because the sale of the first priority is internal to the organization.[30] By comparing (23) and (14), the comparison of investment incentives with vertical integration and without vertical integration depends on the relative magnitude of $2(m_2 + \theta\Delta_m)$ and m_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, precisely, $\theta > \theta_I \equiv \frac{1}{2} - \frac{m_2}{2\Delta_m}$. 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 $\theta < \theta_I$, an independent ISP has greater incentive to invest than a vertically merged one. Note that this condition is identical to the one ensuring 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 the 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.

Vertical integration can also alleviate the holdup problem under the discriminatory regime. It thus could be an alternative way to solve the holdup problem if the ISP cannot commit to net neutrality. When θ is high, vertical integration increases both the ISP's and vertically integrated CP's investments.

8. Concluding remarks

■ This article provides an economic analysis of net neutrality regulation. In particular, our analysis focuses on the effects of net neutrality regulation on the investment incentives of Internet secrvice providers and content providers 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 multitiered network vis-à-vis in a nondiscriminatory network under net neutrality regulation depends on a potential tradeoff between the two sides of the market: the network access fee from 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 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 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 (see Armstrong, 2006 and Rochet and Tirole, 2006 for details). 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

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

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 and helps delivery speed of other content providers. Net neutrality regulation may have a differential effect across different types of investments and impact the choice of investment.

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 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 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.

Appendix A

□ Proofs for lemmas and propositions

Proof of Lemma 1. 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 the nonpriority CP is greater. As a result, we may end up with a situation in which all consumers subscribe to the CP with the first priority. Ironically, in this outcome no one has priority because everyone is treated equally within the priority class. To prevent this outcome from prevailing, we require the two CPs to be sufficiently differentiated. An interior solution with market sharing requires $\widetilde{x} < 1$, that is,

$$v - a - \Gamma_1(\widetilde{x} = 1) < v - a - \Gamma_2(\widetilde{x} = 1), \tag{A1}$$

where $\Gamma_1(\widetilde{x}) = \frac{1}{\mu - \widetilde{x}\lambda} + t\widetilde{x}$ and $\Gamma_2(\widetilde{x}) = \frac{\mu}{\mu-\lambda}\frac{1}{\mu-\widetilde{x}\lambda} + t(1-\widetilde{x})$. This condition is satisfied if $t > \frac{\lambda}{(\mu-\lambda)^2}$.

Moreover, for the adjustment process to yield a stable (interior) equilibrium, we need to assume

$$\Gamma_1'(\widetilde{x}) > \Gamma_2'(\widetilde{x}) \text{ for all } \widetilde{x} \in [1/2, 1]. \tag{A2}$$

By taking the derivatives of $\Gamma_1(\widetilde{x})$ and $\Gamma_2(\widetilde{x})$, we can explicitly write condition (A2) as

$$\frac{\lambda^2}{(\mu - \lambda)(\mu - \widetilde{x}\lambda)^2} < 2t. \tag{A3}$$

Because the left-hand side of (A3) can be alternatively expressed as

$$\frac{\lambda^2}{(\mu - \lambda)(\mu - \widetilde{x}\lambda)^2} = \frac{\lambda}{(\mu - \widetilde{x}\lambda)}\frac{\lambda}{(\mu - \lambda)(\mu - \widetilde{x}\lambda)} = \frac{\lambda}{(\mu - \widetilde{x}\lambda)}(2\widetilde{x} - 1)t, \tag{A4}$$

where the last equality comes from equation (7) which defines $\widetilde{x}$, the stability condition holds if the inequality $\frac{\lambda}{(\mu-\widetilde{x}\lambda)}(2\widetilde{x} - 1) < 2$ is met. Because $\frac{\lambda}{(\mu-\widetilde{x}\lambda)}(2\widetilde{x} - 1)$ is increasing in $\widetilde{x}$ of which the maximum value is one, it can easily be seen that the above inequality is satisfied even for $\widetilde{x} = 1$, if $\mu > \frac{3}{2}\lambda$. On the other hand, by totally differentiating (7), we find the following relationship of

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

Hence, with the assumption of $\mu > \frac{3\lambda}{2}$, we have $\frac{d\widetilde{x}}{d\mu} < 0$ from (A3).

Proof of Lemma 2. 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 between network access fees is given by

$$a^* - \widetilde{a} = \frac{1}{\mu - \widetilde{x}\lambda} - \frac{1}{\mu - \lambda} + t(\widetilde{x} - \frac{1}{2}). \tag{A5}$$

Because $\widetilde{x}$ defined by (7) satisfies the equality of

$$t(2\widetilde{x} - 1) = \frac{1}{\mu - \widetilde{x}\lambda}\frac{\mu}{\mu - \lambda}, \tag{A6}$$

we can prove the given result by dividing (A6) by 2 and substituting (A6) into (A5) as

$$a^* - \widetilde{a} = \frac{1}{\mu - \widetilde{x}\lambda} - \frac{1}{\mu - \lambda} + \frac{1}{2}\frac{1}{\mu - \widetilde{x}\lambda}\frac{\mu}{\mu - \lambda} = \frac{(2\widetilde{x} - 1)\lambda}{2(\mu - \widetilde{x}\lambda)(\mu - \lambda)} > 0 \because \widetilde{x} > 1/2.$$

Proof of Lemma 3. (i) Let $\widetilde{M}$ and M denote the total margin in a discriminatory and in the neutral network, respectively. Then, the difference of these two is

$$\Delta M \equiv \widetilde{M} - M = (\widetilde{x}m_1 + (1-\widetilde{x})m_2)\lambda - \frac{m_1+m_2}{2}\lambda$$
$$= \left(\widetilde{x} - \frac{1}{2}\right)\Delta_m\lambda \geq 0 \ \because \widetilde{x} > 1/2 \text{ and } \Delta_m \geq 0.$$

(ii) Let $\widetilde{T}$ and T denote transaction cost in a discriminatory and in the neutral network, respectively. Then, it is easily seen that

$$\Delta T \equiv \widetilde{T} - T = \left(\int_0^{\widetilde{x}} tx\,dx + \int_{\widetilde{x}}^1 t(1-x)\,dx\right) - \frac{t}{4} = \left(\widetilde{x} - \frac{1}{2}\right)^2 t \geq 0.$$

(iii) Let $\widetilde{W}$ and W denote the total waiting cost in a discriminatory and in the neutral network, respectively. Recall that the expected waiting cost for each end user in the neutral network is given by $w = \frac{1}{\mu-\lambda}$. With the total mass of end users normalized to one, we have that $W = w = \frac{1}{\mu-\lambda}$. By contrast, $\widetilde{W}$ is calculated as the weighted average of $w_1 = \frac{1}{\mu-\lambda_1}$ and $w_2 = \frac{\mu}{\mu-\lambda}w_1$, with weights $\widetilde{x}$ and $(1-\widetilde{x})$, respectively. Using $\lambda_1 = \widetilde{x}\lambda$, we find that

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

Proof of Lemma 4. Consider an equilibrium in which CP1 has a higher *ex post* margin (that is, $\bar{m}_1 + \Delta_1 > \bar{m}_2 + \Delta_2$) and receives the priority. In such a case, CP1 solves the following problem:

$$\max_{\Delta_1}(\bar{m}_1 + \Delta_1)\widetilde{x}\lambda - [(\bar{m}_2 + \Delta_2) + \theta((\bar{m}_1 + \Delta_1) - (\bar{m}_2 + \Delta_2))](2\widetilde{x} - 1)\lambda - \Psi(\Delta_1).$$

Then, it is clear that CP1's optimal investment level for this case is given by $\bar{\Delta}$ by the first-order condition. Similarly, CP2's optimal investment problem is given by

$$\max_{\Delta_2}(\bar{m}_2 + \Delta_2)(1-\widetilde{x})\lambda - \Psi(\Delta_2).$$

Thus, CP2's optimal investment level is given by $\underline{\Delta}$. We have derived these equilibrium investment levels assuming that CP1 receives the priority.

For the investment levels $(\bar{\Delta}, \underline{\Delta})$ to be sustainable as an equilibrium, we need to check whether CP2 will have an incentive to invest up to the level where it will be the firm that receives the priority with a higher *ex post* margin, given that CP1 invests at $\bar{\Delta}$. This deviation requires an investment level of at least $\bar{\Delta} + (\bar{m}_1 - \bar{m}_2)$ by CP2. Because we know that $\bar{\Delta}$ is the optimal investment level for the firm that acquires the priority right and the objective function is concave, the optimal deviation investment level is given by $\bar{\Delta} + (\bar{m}_1 - \bar{m}_2)$. The optimal deviation payoff for CP2 is thus given by

$$\widetilde{\pi}_2^d = (\bar{m}_1 + \bar{\Delta})\widetilde{x}\lambda - (\bar{m}_1 + \bar{\Delta})(2\widetilde{x} - 1)\lambda - \Psi(\bar{\Delta} + (\bar{m}_1 - \bar{m}_2)).$$

However, we have

$$\widetilde{\pi}_2^d = (\bar{m}_1 + \bar{\Delta})(1-\widetilde{x})\lambda - \Psi(\bar{\Delta} + (\bar{m}_1 - \bar{m}_2))$$
$$= [\bar{m}_2 + (\bar{\Delta} + (\bar{m}_1 - \bar{m}_2))](1-\widetilde{x})\lambda - \Psi(\bar{\Delta} + (\bar{m}_1 - \bar{m}_2))$$
$$= \max_{\Delta_2}(\bar{m}_2 + \Delta_2)(1-\widetilde{x})\lambda - \Psi(\Delta_2)$$
$$= (\bar{m}_2 + \underline{\Delta})(1-\widetilde{x})\lambda - \Psi(\underline{\Delta}).$$

This implies that CP2 has no incentive to deviate from investing $\underline{\Delta}$.

Finally, we check whether CP1 will have an incentive to deviate by reducing its investment to the level where it ends up at a lower *ex post* margin given that the CP2 invests at $\underline{\Delta}$, which is possible if $\bar{m}_1 < \bar{m}_2 + \underline{\Delta}$. By following the logic above, the optimal deviation investment level is given by $\underline{\Delta} - (\bar{m}_1 - \bar{m}_2)$ because $\underline{\Delta}$ is the optimal investment level for the firm that fails to receive the priority right and the objective function is concave. The optimal deviation payoff for CP1 is given by

$$\widetilde{\pi}_1^d = [\bar{m}_1 + (\underline{\Delta} - (\bar{m}_1 - \bar{m}_2))](1-\widetilde{x})\lambda - \Psi(\underline{\Delta} - (\bar{m}_1 - \bar{m}_2)).$$

Let $\bar{\Delta}(\theta)$ and $\widetilde{\pi}_1^*(\theta)$ denote CP1's investment level and its corresponding payoff in the putative equilibrium. Note that $\bar{\Delta}(\theta)$ is decreasing in θ and $\bar{\Delta}(\theta = 1) = \underline{\Delta}$. By the envelope theorem, we also know that $\widetilde{\pi}_2^*(\theta)$ is a decreasing function of θ. Thus, for all $\theta \in [0, 1]$, we have

$$\begin{aligned}
\widetilde{\pi}_1^*(\theta) &\geq \widetilde{\pi}_1^*(\theta = 1) \\
&= [\bar{m}_1 + \bar{\Delta}(\theta = 1)]\widetilde{x}\lambda - [(\bar{m}_2 + \underline{\Delta}) + 1 \cdot ((\bar{m}_1 + \bar{\Delta}(\theta = 1)) - (\bar{m}_2 + \underline{\Delta}))](2\widetilde{x} - 1)\lambda - \Psi(\bar{\Delta}(\theta = 1)) \\
&= [\bar{m}_1 + \bar{\Delta}(\theta = 1)](1 - \widetilde{x})\lambda - \Psi(\bar{\Delta}(\theta = 1)) \\
&= (\bar{m}_1 + \underline{\Delta})(1 - \widetilde{x})\lambda - \Psi(\underline{\Delta}) \\
&> [\bar{m}_1 + (\underline{\Delta} - (\bar{m}_1 - \bar{m}_2))](1 - \widetilde{x})\lambda - \Psi(\underline{\Delta} - (\bar{m}_1 - \bar{m}_2)) = \widetilde{\pi}_1^d.
\end{aligned}$$

The last inequality above comes from the definition of $\underline{\Delta}$ and the preceding equality comes from the fact that $\bar{\Delta}(\theta = 1) = \underline{\Delta}$. Thus, CP1 does not have an incentive to deviate, either. Taken together, we can conclude that $(\bar{\Delta}, \ \underline{\Delta})$ is an equilibrium. This is the unique equilibrium if the margin difference $(\bar{m}_1 - \bar{m}_2)$ is sufficiently large. Otherwise, we can have another pure strategy equilibrium in which the roles are reversed. We can proceed in a similar way to show that $(\underline{\Delta}, \ \bar{\Delta})$ can also be an equilibrium if $(\bar{m}_1 - \bar{m}_2)$ is sufficiently small.

Proof of Proposition 1. Recall that we have derived the ISP's profits under a neutral regime and discriminatory regime as $\pi_m^* = a^* = v - \frac{1}{\mu-\lambda} - \frac{t}{2}$ and $\widetilde{\pi}_m^* = (v - \frac{1}{\mu-\widetilde{x}\lambda} - t\widetilde{x}) + (m_2 + \theta\Delta_m)(2\widetilde{x} - 1)\lambda$, respectively. The difference between these two is given by

$$\begin{aligned}
\widetilde{\pi}_m^* - \pi_m^* &= \left(v - \frac{1}{\mu - \widetilde{x}\lambda} - t\widetilde{x}\right) + (m_2 + \theta\Delta_m)(2\widetilde{x} - 1)\lambda - \left(v - \frac{1}{\mu - \lambda} - \frac{t}{2}\right) \\
&= \left(\frac{1}{\mu - \lambda} - \frac{1}{\mu - \widetilde{x}\lambda}\right) + (2\widetilde{x} - 1)\left\{(m_2 + \theta\Delta_m)\lambda - \frac{t}{2}\right\} > 0 \\
&\iff (m_2 + \theta\Delta_m) > \Lambda \equiv \frac{t}{2\lambda} + \frac{1}{(2\widetilde{x} - 1)\lambda}\left(\frac{1}{\mu - \widetilde{x}\lambda} - \frac{1}{\mu - \lambda}\right).
\end{aligned}$$

Hence, if $m_2 > \Lambda$, $\widetilde{\pi}_m^* > \pi_m^*$ for all $\theta \in [0, 1]$, which proves the statement (i). Conversely, because the term $(m_2 + \theta\Delta_m)$ is increasing in θ, we get $\widetilde{\pi}_m^* < \pi_m^*$ for all $\theta \in [0, 1]$ when $m_2 < \Lambda$. Finally, if $m_2 \leq \Lambda \leq m_1$, there exists a critical level of $\theta^* \in [0, 1]$ such that $\widetilde{\pi}_m^* > \pi_m^*$ iff $\theta > \theta^*$.

Proof of Proposition 2. The statements in (i) and (ii) can be proved in a straightforward manner by comparing the expressions for the CPs' profits across the regimes. Concerning the statement in (iii), let us denote the aggregate consumer welfare in 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 $CS = 2\int_0^{1/2} tx\,dx$ and $\widetilde{CS} = \int_0^{\widetilde{x}} tx\,dx + \int_0^{1-\widetilde{x}} tx\,dx$. Aggregate consumer welfare increases in the discriminatory network, which is simply shown as

$$\widetilde{CS} - CS = \int_0^{\widetilde{x}} tx\,dx + \int_0^{1-\widetilde{x}} tx\,dx - 2\int_0^{1/2} tx\,dx = \int_{1/2}^{\widetilde{x}} tx\,dx - \int_{1-\widetilde{x}}^{1/2} tx\,dx > 0.$$

Proof of Proposition 3. The comparison of social welfare across the two different regimes can be readily seen by the sign of $\Delta M - \Delta T$.

$$\begin{aligned}
\Delta M - \Delta T &\equiv \widetilde{M} - M - (\widetilde{T} - T) = \left(\widetilde{x} - \frac{1}{2}\right)(m_1 - m_2)\lambda - \left(\widetilde{x} - \frac{1}{2}\right)^2 t \\
&= \left(\widetilde{x} - \frac{1}{2}\right)\left\{\Delta_m\lambda - \left(\widetilde{x} - \frac{1}{2}\right)t\right\} > 0 \\
&\iff \Delta_m\lambda - \left(\widetilde{x} - \frac{1}{2}\right)t > 0 \iff \Delta_m > \left(\widetilde{x} - \frac{1}{2}\right)\frac{t}{\lambda}.
\end{aligned}$$

Thus, if the margin difference is significantly large compared to the degree of product differentiation to the extent of $\Delta_m > (\widetilde{x} - \frac{1}{2})\frac{t}{\lambda}$, the social welfare is higher even in the discriminatory network, precisely.

Proof of Proposition 5. First, let us recall that the marginal changes in the CPs' profits are given by $\Psi'(\Delta_i^*) = \frac{\lambda}{2}$ and $\Psi'(\widetilde{\Delta}_1^*) = \{\widetilde{x} - \theta(2\widetilde{x} - 1)\}\lambda$. With the assumption of $\Psi'' > 0$, we get $\Delta_1^* > \widetilde{\Delta}_1^*$ if and only if $\frac{1}{2} > \widetilde{x} - \theta(2\widetilde{x} - 1)$. From the following relationship,

$$\begin{aligned}
\frac{1}{2} - \{\widetilde{x} - \theta(2\widetilde{x} - 1)\} &= (2\widetilde{x} - 1)\left(\theta - \frac{1}{2}\right) > 0 \\
&\iff \theta > \tfrac{1}{2}\ (\because \widetilde{x} > 1/2),
\end{aligned}$$

we derive the result (i) that $\theta > \frac{1}{2}$ is the necessary and sufficient condition for $\Delta_i^* > \widetilde{\Delta}_1^*$. For $\theta \in [0, 1/2]$, we have the opposite case of $\widetilde{\Delta}_1^* \geq \Delta_1^*$. Similarly, the comparison between $\Psi'(\Delta_2^*) = \frac{\lambda}{2}$ and $\Psi'(\widetilde{\Delta}_2^*) = (1 - \widetilde{x})\lambda$ yields the result of $\widetilde{\Delta}_2^* < \Delta_2^*$ for $\forall\theta \in [0, 1]$.

Proof of Proposition 7. Recall that $\widetilde{\theta}$ must be set such that $\frac{d\widetilde{\pi}_m^*}{d\theta} = \frac{\partial\widetilde{\pi}_m^*}{\partial\theta} + \frac{\partial\widetilde{\pi}_m^*}{\partial\widetilde{\Delta}_1^*}\frac{\partial\widetilde{\Delta}_1^*}{\partial\theta} = 0$. Because we can easily obtain both $\frac{\partial\widetilde{\pi}_m^*}{\partial\theta} = \Delta_m(2\widetilde{x} - 1)\lambda$ and $\frac{\partial\widetilde{\pi}_m^*}{\partial\widetilde{\Delta}_1^*} = \theta(2\widetilde{x} - 1)\lambda$ from (11) as well as $\frac{\partial\widetilde{\Delta}_1^*}{\partial\theta} = -(2\widetilde{x} - 1)k\lambda$ from $\Psi'(\widetilde{\Delta}_1^*) = \{\widetilde{x} - \theta(2\widetilde{x} - 1)\}\lambda$ with $\Psi(\Delta_i) = \Delta_i^2/2k$, the optimal level of rent extraction is derived as

$$\begin{aligned}\frac{d\widetilde{\pi}_m^*}{d\theta} &= \Delta_m(2\widetilde{x} - 1)\lambda - \theta(2\widetilde{x} - 1)^2 k\lambda^2 \\ &= (2\widetilde{x} - 1)\lambda\{\Delta_m - \theta(2\widetilde{x} - 1)k\lambda\} = 0 \text{ at } \theta = \widetilde{\theta}.\end{aligned}$$

Hence, we get $\widetilde{\theta} = \frac{\Delta_m}{(2\widetilde{x}-1)k\lambda}$. The results of $\frac{\partial\widetilde{\theta}}{\partial k} < 0$ and $\frac{\partial\widetilde{\theta}}{\partial\Delta_m} > 0$ are immediate from the comparative statics for $\widetilde{\theta}$ with respect to k and Δ_m, respectively.

Appendix B

□ **Exclusive versus nonexclusive priority.** CP_i's willingness to pay for the exclusive right to the fast lane will depend upon whether CP_j, for $i, j \in \{1, 2\}$ and $j \neq i$, will be granted that right if CP_i does not buy it. In this appendix, we derive conditions under which selling an exclusive right is preferred to the selling of a nonexclusive right. For this purpose, we assume that the ISP has all the bargaining power and has the ability to make take-it-or-leave-it offers to CPs, as in Armstrong (1999). Let us denote by b_i and l_i the benefits and losses associated with the exclusive right, respectively. The net gain to accepting the exclusive right is the sum of b_i and l_i: CP_i will pay up to

$$b_i + l_i = \left(\widetilde{x}\lambda m_i - \frac{\lambda}{2}m_i\right) + \left\{\frac{\lambda}{2}m_i - (1 - \widetilde{x})\lambda m_i\right\} = (2\widetilde{x} - 1)\lambda m_i \tag{B1}$$

for the exclusive right. Therefore, the highest rent that the ISP can obtain by selling the exclusive right to first priority is to sell to the firm with the competitive advantage in margin (CP1), in which case its rent is

$$R_{ex} = (2\widetilde{x} - 1)\lambda m_1. \tag{B2}$$

In the discriminatory regime, the ISP can obtain a total profit of[31]

$$\widetilde{a} + R_{ex} = \left(v - \frac{1}{\mu - \widetilde{x}\lambda} - t\widetilde{x}\right) + (2\widetilde{x} - 1)\lambda m_i. \tag{B3}$$

Now let us explore whether it is optimal for the ISP to sell the right to first priority exclusively to only one CP. Suppose that the ISP offers to sell the right to CP_i for a charge R_i and that both CPs accept. Then in order for i to agree to pay this charge (given that j has also agreed), we must have R_i no greater than the profit loss from not having the right when firm j does, which is just l_i. Therefore, the most the ISP can get from selling the nonexclusive right is just $l_1 + l_2$.[32] In this case, the ISP will obtain a profit

$$\begin{aligned}a^* + R_{non} &= \left(v - \frac{1}{\mu - \lambda} - \frac{t}{2}\right) + \sum_{i=1}^{2}\left(\frac{\lambda}{2}m_i - (1 - \widetilde{x})\lambda m_i\right) \\ &= \left(v - \frac{1}{\mu - \lambda} - \frac{t}{2}\right) + (2\widetilde{x} - 1)\lambda\frac{(m_1 + m_2)}{2}.\end{aligned} \tag{B4}$$

Therefore, the ISP prefers selling the right exclusively to one CP if and only if

$$(\widetilde{a} + R_{ex}) - (a^* + R_{non}) > 0 \Longleftrightarrow R_{ex} - R_{non} > a^* - \widetilde{a}.$$

With simple algebra, we know that $R_{ex} - R_{non} = (2\widetilde{x} - 1)\lambda\frac{\Delta_m}{2}$ and $a^* - \widetilde{a} = \frac{(2\widetilde{x}-1)\lambda}{2(\mu-\widetilde{x}\lambda)(\mu-\lambda)}$. A simple manipulation yields that the ISP will use the exclusive scheme if the margin differential is sufficiently large to the extent of

$$\Delta_m > \frac{1}{(\mu - \widetilde{x}\lambda)(\mu - \lambda)}. \tag{B5}$$

[31] It is important that the ISP may be able to commit to granting the exclusive right to CP2 in the event that CP1 rejects the ISP's offer, if it is in the interests of the ISP to do so. A simple scheme that appears to avoid some of these commitment issues is to auction off the exclusive right to the highest bidder. In this case, the bidding will stop when CP2 drops out at the price $b_2 + l_2$.

[32] This assumes that the ISP can make discriminatory offers to CPs. Note that Cheng et al. (2009), by contrast, assume that the ISP makes nondiscriminatory offers. With nondiscriminatory offers, an exclusive contract would be more attractive to the ISP.

References

ANDERSON, S.P. AND COATE, S. "Market Provision of Broadcasting: A Welfare Analysis." *Review of Economic Studies,* Vol. 72 (2005), pp. 947–972.

ARMSTRONG, M. "Competition in the Pay-TV Market." *Journal of the Japanese and International Economies,* Vol. 13 (1999), pp. 257–280.

——. "Competition in Two-Sided Markets." *RAND Journal of Economics,* Vol. 37 (2006), pp. 668–691.

BALACHANDRAN, K.R. "Purchasing Priorities in Queues." *Management Science,* Vol. 18 (1972), pp. 319–326.

BUSINESS WEEK ONLINE. "Rewired and Ready for Combat." November 7, 2005. Available at www.businessweek.com/magazine/content/05_45/b3958089.htm (accessed on June 29, 2010).

CAILLAUD, B. AND JULLIEN, B. "Chicken and Egg: Competition among Intermediation Service Providers." *RAND Journal of Economics,* Vol. 34 (2003), pp. 309–328.

CERF, V.G. "Prepared Statement in the U.S. Senate Committee on Commerce, Science, and Transportation Hearing on 'Net Neutrality.' " February 7, 2006.

CHENG, H.K., BANDYOPADHYAY, S., AND GUO, H. "The Debate on Net Neutrality: A Policy Perspective." *Information Systems Research*, (2009).

CHIRICO, F., VAN DE HAAR, I., AND LAROUCHE, P. "Network Neutrality in the EU." Discussion Paper no. DP 2007-030, TILEC, 2007.

CHOI, J.P. "Optimal Tariffs and the Choice of Technology: Discriminatory Tariffs vs. the 'Most Favored Nation' Clause." *Journal of International Economics,* Vol. 58 (2010), pp. 560–579.

——. "Tying in Two-Sided Markets with Multi-Homing." *Journal of Industrial Economics*.

CREMER, J., REY, P., AND TIROLE, J. "Connectivity in the Commercial Internet." *Journal of Industrial Economics,* Vol. 48 (2000), pp. 433–472.

DeGraba, P. "Input Market Price Discrimination and the Choice of Technology." *American Economic Review*, Vol. 80 (1990), pp. 1246–1253.

DENECKERE, R.J. AND MCAFEE, P. "Damaged Goods." *Journal of Economics & Management Strategy,* Vol. 5 (1996), pp. 149–174.

ECONOMIDES, N. "Net Neutrality, Non-Discrimination and Digital Distribution of Content through the Internet." *A Journal of Law and Policy for the Information Society,* Vol. 4 (2008), pp. 209–233.

—— AND TÅG, J. "Net Neutrality on the Internet: A Two-Sided Market Analysis." Working Paper no. 07-14, NET Institute, New York University, 2007.

EDELSON, N.M. AND HILDERBRAND, D.K. "Congestion Tolls for Poisson Queuing Processes." *Econometrica,* Vol. 43 (1975), pp. 81–92.

FUDENBERG, D. AND TIROLE, J. "Understanding Rent Dissipation: On the Use of Game Theory in Industrial Organization." *American Economic Review: Papers and Proceedings,* Vol. 77 (1987), pp. 176–183.

GROSS, D. AND HARRIS, C.M. *Fundamentals of Queueing Theory*, 3rd ed. New York: Wiley, 1998.

HAHN, J.H. "Damaged Durable Goods." *RAND Journal of Economics,* Vol. 37 (2006), pp. 121–133.

HAHN, R.W. AND WALLSTEN, S. "The Economics of Net Neutrality." Working Paper no. RP06-13, AEI-Brookings Joint Center, 2006. Available at ssrn.com/abstract=943757 (accessed June 29, 2010).

HERMALIN, B.E. AND KATZ, M.L. "The Economics of Product-Line Restrictions with an Application to the Network Neutrality Debate." *Information Economics and Policy,* Vol. 19 (2007), pp. 215–248.

KOCSIS, V. AND DE BIJL, P. "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.

MENDELSON, H. "Pricing Computer Services: Queueing Effects." *Communications of the ACM,* Vol. 28 (1985), pp. 312–321.

—— AND WHANG, S. "Optimal Incentive-Compatible Priority Pricing for the M/M/1 Queue." *Operations Research,* Vol. 38 (1990), pp. 870–883.

NAOR, P. "The Regulation of Queue Size by Levying Tolls." *Econometrica,* Vol. 37 (1969), pp. 15–24.

ROCHET, J.C. AND TIROLE, J. "Two-Sided Markets: A Progress Report." *RAND Journal of Economics,* Vol. 37 (2006), pp. 645–667.

VALLETTI, T. AND CAMBINI, C. "Investments and Network Competition." *RAND Journal of Economics,* Vol. 36 (2005), pp. 446–468.

VAN SCHEWICK, B. "Towards an Economic Framework for Network Neutrality Regulation." *Journal on Telecommunications and High Technology Law,* Vol. 5 (2007), pp. 329–392.

VON HIPPEL, E. *Democratizing Innovation*. Cambridge, Mass.: MIT Press, 2005.

WU, T. "Network Neutrality, Broadband Discrimination." *Journal on Telecommunications and High Technology Law,* Vol. 2 (2003), pp. 141–178.

YOO, C.S. "Network Neutrality and the Economics of Congestion." *Georgetown Law Journal,* Vol. 94 (2006), pp. 1847–1908.