



**HAL**  
open science

## Zero-rating, content quality and network capacity

Emmanuel Lorenzon

► **To cite this version:**

Emmanuel Lorenzon. Zero-rating, content quality and network capacity. Information Economics and Policy, 2022, 58, 10.1016/j.infoecopol.2022.100965 . hal-03558783

**HAL Id: hal-03558783**

**<https://hal.science/hal-03558783>**

Submitted on 22 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Zero-Rating, Content Quality, and Network Capacity

Emmanuel Lorenzon<sup>1,2</sup>

<sup>1</sup>Univ. Grenoble Alpes, CNRS, INRAE, Grenoble INP, GAEL, 38000 Grenoble, France

<sup>2</sup>Governance and Regulation Chair, University Paris-Dauphine, PSL Research University, 75016 Paris, France

January 20, 2022

## Abstract

We consider a departure from net neutrality by an Internet service provider (ISP) that financially discriminates among content providers through exclusive zero-rating contracts. Zero-rating is an instrument to distort competition between content providers and the manner in which consumers value content. We analyze its implications for the incentives to provide quality in the market for content and to invest in broadband infrastructure. Zero-rating makes content more expensive for consumers to use and imply a downward distortion of content quality. Content providers switch from minimal differentiation to a downward vertical differentiation outcome. Next, we find that zero-rating implies under-provision in the broadband infrastructure, which comes from a standard rent-extraction argument and a cost-alleviation channel related to the complementarity between network capacity and content quality. Finally, when implemented, zero-rating is found to be welfare reducing and detrimental to consumers.

*Keywords:* Internet; Net neutrality; Zero-rating; Network capacity; Content quality; Congestion; Three-part tariff

*JEL Classification:* D21, L12, L51, L96, R41

Declarations of interest: none

# 1 Introduction

Net neutrality, according to which an Internet service provider (ISP) should not discriminate among data packets sent on its network, has been a matter of heated debate during the past decade, with new developments in recent years. In Europe, the practice of zero-rating is pervasive (European Commission, 2017), but related legislation is still debated and opaque<sup>1</sup>. In September 2020, The Court of Justice of the European Union enshrined the net neutrality principle by ruling against a zero-rating practice used by the Hungarian telecom operator Telenor, which has been found to use zero-rating as a traffic management tool to discriminate against applications that were not part of its MyChat and MyMusic subscription plans<sup>2</sup>.

Zero-rating is a practice by which an ISP makes some content more expensive than others for consumers to access. Consumers subscribe to a monthly mobile data plan, which provides a data allowance. For all data packets consumed in excess of that allowance, consumers are charged marginal fees and/or the data usage is either blocked or restricted. Thus, zero-rating is a tool that an ISP can implement to price discriminate among content providers (CPs). The data from zero-rated content do not count against the cap. Once a consumer reaches his or her data cap, such content is exempted from per-unit surcharges and usage restrictions<sup>3</sup>.

Network operators advertise that zero-rating is beneficial for users because it purportedly allows users to consume more content at the same price for their mobile plan and allows operators to efficiently manage traffic and foster their incentives to invest in network quality (Schnurr and Wiewiorra, 2018, and Krämer and Peitz, 2018). Because CPs typically rely on traffic to generate revenue from either advertisements or user payments, opponents of such data management regimes contend that departures from net neutrality might steer consumers' choices of online content toward providers included in the operator's contracting offers, which could raise barriers to entry and impede incentives to provide high content quality in the market for content (Nurski, 2012). Although the issues surrounding investments by ISPs in network infrastructure are crucial for regulators and well documented, especially with respect to priority pricing, their interplay with content quality has yet to be overlooked by the network neutrality literature regarding zero-rating practices (Krämer and Peitz, 2018). However, a tight relationship exists between the two dimensions. For instance, incentives to provide qualitative network infrastructure are positively related to consumers' willingness to pay for content, which itself positively depends on the quality of the content provided within the network.

We add to the growing debate on data management regimes by focusing on investment decisions made by CPs and a profit-maximizing ISP. As in Gautier and Somogyi (2020), zero-

---

<sup>1</sup>The report is available at <https://ec.europa.eu/competition/publications/reports/kd0217687enn.pdf>. In some countries, regulators punish operators offering zero-rating plans to their consumers. For example, Sweden's telecom regulator, PTS, has ordered the operator Telia to stop selling contracts with unlimited data for selected social media and streaming services because they are not compliant with the net neutrality rules under the EU's Telecom Single Market regulation. ("Net Neutrality's Holes in Europe May Offer Peek at Future in U.S.," New York Times, December 10, 2017. Available at <https://www.nytimes.com/2017/12/10/business/net-neutrality-europe-fcc.html>).

<sup>2</sup>The judgement of the court is available at this [link](#).

<sup>3</sup>In contrast, paid prioritization is a practice that also violates net neutrality and, according to which, in exchange for side payments from CPs, an ISP creates a "fast lane" to prioritize the delivery time of certain content over others.

rating is viewed as a tool to enhance content differentiation and to alter the competition in the market for content. However, the focus has been on the ability of an ISP to mitigate content asymmetry in environments in which CPs are passive and network investments are left separate<sup>4</sup>. CPs do not invest in content attractiveness or determine the amount of advertisement to which to expose their users, and ISPs face some given capacity constraints. Our aim is to study the implications that departure from a network neutrality regulation has for ISPs and CPs' incentives to provide broadband facilities and content quality.

We compare two regulatory regimes: net neutrality and bilateral zero-rating contracts, which means that the ISP offers an exclusive zero-rating contract to a unique CP<sup>5</sup>. To study the impact of a move from the net neutrality regime on the investment decisions of CPs and the ISP, we consider a monopolistic ISP connecting two horizontally differentiated CPs to a unit mass of consumers with unit demand, which is consistent with the fact that users typically choose one CP at a time for each device or ISP. The ISP provides a network capacity constraint, and CPs compete *à la* Hotelling to offer vertically differentiated services to consumers under an advertising-supported service model. The ISP charges consumers a connection fee to access its network and might price discriminate between CPs by charging consumers different per-unit fees for the two content offerings. CPs differ in their advertising revenues, and each imposes different advertising exposure levels on consumers on their respective websites. Thus, horizontal and vertical differentiation interact, and CPs compete on the level of advertisement to which they expose their users and from which they draw their revenues. Finally, we suppose that CPs are asymmetric with respect to the number of requests that users generate for their respective content, in that one CP obtains more content requests than its competitor.

In our model, "quality" is understood as investments by CPs in the attributes or functionalities embodied in their content. In addition, consumers bear disutility per unit of content from being exposed to advertisements that, as in [Calzada and Tselekounis \(2018\)](#), interact with content quality. The idea is that higher content quality means that consumers spend more time using it and are more exposed to advertisements, and CPs benefit more from quality improvements. Hence, the per-unit advertising exposure rate is understood as a per-unit fee that CPs charge consumers, meaning that it possesses the same properties as a pure per-unit price to access the content. As a result, our setup not only allows for the investigation of quality and advertisement competition between CPs but also captures the business models of most services included in zero-rating<sup>6</sup>.

The main contribution of this work to the literature is to demonstrate that, in a static model with asymmetric CPs, a zero-rating makes content more expensive to access, implying a downward distortion of quality in the market for content and provision of broadband capacity

---

<sup>4</sup>See, for instance, [Gautier and Somogyi \(2020\)](#), [Jeitschko et al. \(2020\)](#), or [Jullien and Sand-Zantman \(2018\)](#).

<sup>5</sup>We consider a strict application of net neutrality rules that forbids the ISP from charging a linear access price to either CPs to obtain access to the network or consumers for any data consumed in excess of their initial allowance. We do not consider other forms of discrimination, such as paid-prioritization contracts.

<sup>6</sup>For example, Netflix generates revenues exclusively from user subscriptions, social media websites such as Facebook and Instagram rely largely on advertising revenue, and streaming media services such as YouTube and Spotify, as well as many other online applications, use a "freemium" model in which revenue is generated from advertising displayed to consumers using the service for free while charging consumers to access advertising-free content.

attributable to complementarity between content quality and network capacity.

Our first result is to show that a departure from net neutrality softens quality competition and constitutes an impediment to quality improvement in the market for content. While in a neutral network, content quality is symmetric and CPs opt for minimal vertical differentiation, a departure from net neutrality reduces the overall level of content quality in the market and increases the degree of asymmetry between CPs. Incentives to invest in content quality are misaligned between CPs. The least attractive CP, which has a zero-rating contract with the ISP, benefits from greater market share and provides higher quality than its competitor to increase users' willingness to switch to its content and reduce the disutility from congestion supported by its users. In contrast, the noncontracting CP opts for quality degradation to reduce the ISP's ability to price discriminate between consumers and to prevent switching by its home users. However, investments provided by the zero-rated CP do not sufficiently compensate for the lack of investment from its rival, making the content sector less innovative. As a result, CPs switch from a minimal vertical differentiation outcome under net neutrality to an asymmetric equilibrium with a greater degree of downward vertical differentiation when zero-rating is allowed. In parallel, we find that zero-rating softens advertising competition in the market for content. Users are more exposed to advertising than in a neutral network. Because advertising exposure can be interpreted as a per-unit payment from consumers to CPs, content becomes more expensive to access when zero-rating is implemented.

Our second main result is to show that a departure from net neutrality reduces network investments. Alongside a price discrimination argument, which implies that the ISP offers a zero-rating contract to the least attractive CP, zero-rating reduces the ISP's incentives to invest in network capacity through a cost-alleviation channel. Indeed, in contrast to the main claim from the industry, we find that a profit-maximizing ISP underinvests under a discriminatory regime. However, given the complementarity with content quality, we demonstrate that zero-rating reduces network congestion, which might constitute a traffic management instrument. This underinvestment is motivated by two explanatory channels. The first operates via standard price discrimination and relates to the rent-extraction effect derived by [Choi and Kim \(2010\)](#): the ISP increases resource scarcity, which allows for a higher per-unit surcharge on consumers of non-zero-rated content. Therefore, in equilibrium, as in [Gautier and Somogyi \(2020\)](#), the ISP offers a zero-rating contract to the least attractive CP. Reducing network capacity allows larger rents to be extracted from users of the most attractive content. The second channel operates via cost-reducing incentives: indirectly, the quality of the non-zero-rated content and network capacity are complements. By imposing a per-unit fee, a strong CP reduces the quality it provides to its users, also reducing congestion on the network and alleviating the need for the ISP to invest in the network; thus, it strategically reduces its investment. Hence, as a contribution to the debate surrounding the efficiency of broadband investment decisions, we find that, in contrast to [Gautier and Somogyi \(2020\)](#), investments in capacity are socially suboptimal under a discriminatory regime. In particular, an ISP underprovides network capacity with respect to the social optimum.

Finally, the social welfare analysis shows that a profit-maximizing ISP fails to adopt a socially efficient pricing policy and zero-rating is always implemented, which is detrimental to consumers and welfare-reducing. The ISP has incentives to engage in practices that make re-

sources scarcer and always finds it profitable to financially discriminate between CPs, whereas consumers are always better off under net neutrality. Hence, a profit-maximizing ISP has pervasive incentives to implement zero-rating contracts, and these incentives are misaligned with consumer welfare.

**Related works** This work contributes to the significant literature on net neutrality and data management practices by ISPs that violate the principle<sup>7</sup>. In particular, this work is close to the body of work that models the impact of net neutrality on ISPs' incentives to invest in network capacity with an emphasis on content innovation<sup>8</sup>.

As underlined by [Goldfarb and Tucker \(2019\)](#), the literature on net neutrality developed along the usage of data transmission technologies and the growing importance of the role of ISP strategies for the emergence of other businesses. [Lee and Wu \(2009\)](#) considered net neutrality as a type of subsidy for innovation to the extent that it does not impose transaction costs on CPs. In their vision, departures from net neutrality oblige CPs to negotiate with ISPs in a similar fashion as in other industries, such as cable TV, potentially discouraging several innovators from creating new services. They contend that in an industry of "stars," such as the CP industry, lower transaction costs are possible because CPs do not need to negotiate with distinct ISPs for access to their consumers and reduce barriers to entry, thereby enabling the emergence of new players. In contrast with this view, and as noted by [Schnurr and Wiewiorra \(2018\)](#), operators advertise paid prioritization and zero-rating as beneficial for users because the latter can consume more content while paying the same price for their mobile plan.

In a deregulated market, ISPs have incentives to depart from net neutrality because they can generate additional revenues from CPs by offering benefits in return (e.g., prioritization of data or exemptions from users' data allowance), attract new customers from the network effects, and better discriminate among consumers on price and quality ([Krämer and Peitz, 2018](#), and [Goldfarb and Tucker, 2019](#)). In [Schnurr and Wiewiorra \(2018\)](#), *laissez-faire* practices might distort competition toward the CPs included in the sponsorship plan and might cause losses to those excluded. The authors consider two symmetric CPs generating revenues from advertising and connected to consumers via a monopolistic ISP. Consumers have preferences for one CP ("high-value") or the other ("low-value"). The authors emphasize that, according to their results, both zero-rating and paid prioritization are distortionary—challenging the view that the latter should be more scrutinized than the former. This result is in line with the empirical findings in [Nurski \(2012\)](#) using data from the United Kingdom. The author finds that departures from net neutrality might steer consumers' choice toward the CP included in the zero-rating plan. Once CPs typically rely on traffic-based revenues (advertising or consumer payment), zero-rating raises barriers to entry for CPs excluded from the plan and might reduce the variety of CPs available to users and the quality provided in the market for content. We contribute to this debate in line with the last observation because we find that a discriminatory regime reduces content quality and implies downward vertical differentiation. Given

---

<sup>7</sup>See [Easley et al. \(2018\)](#) and [Greenstein et al. \(2016\)](#) for a survey.

<sup>8</sup>See, for instance, [Hermalin and Katz \(2007\)](#), [Choi and Kim \(2010\)](#), [Economides and Hermalin \(2012\)](#), [Economides and Tåg \(2012\)](#), [Krämer and Wiewiorra \(2012\)](#), [Bourreau et al. \(2015\)](#), or [Peitz and Schuett \(2016\)](#) for the implications for incentives to invest in broadband capacity and [Krämer and Wiewiorra \(2012\)](#), [Reggiani and Valletti \(2016\)](#), or [Choi et al. \(2018\)](#) for the interplay with content innovation.

the interplay among content quality, congestion, and network capacity, zero-rating reduces managerial costs, increases CPs' asymmetry, and allows the ISP to collect higher payments from users of the non-zero-rated CP.

One important aspect in this setting is that *laissez-faire* data regimes might lead to excess data consumption, resulting in negative externalities in the form of congestion. [Bourreau et al. \(2015\)](#) noted that different services require distinct network capacities and that departing from net neutrality allows ISPs to manage data packages according to content, which in turn enables them to alleviate capacity constraints and congestion effects.

Our representation of zero-rating contract is close to that of [Gautier and Somogyi \(2020\)](#), and our modeling assumption about network investment decisions borrows from [Choi and Kim \(2010\)](#) and [Bourreau et al. \(2015\)](#). As in [Gautier and Somogyi \(2020\)](#), we find that the ISP always price discriminates and contracts with the least attractive CP in equilibrium. Although in their model zero-rating is implemented to reduce the asymmetry between CPs, in our model, the ISP uses zero-rating to distort vertical differentiation between content (which increases CPs' asymmetry) and to affect how consumers value each CP. However, in contrast to their result that investment in capacity is aligned with the social optimum, this is never the case in our model. Private and public incentives to invest are aligned solely under net neutrality regulation, and a discriminatory regime reduces network capacity, which is in line with [Choi and Kim \(2010\)](#), who show that network investments are reduced in the long run<sup>9</sup>. Although for them, this strategic reduction stems from a rent-extraction effect (also presented in our model) arising from congestion, the reduction in capacity in our model follows from the fact that managerial costs are borne by the noncontracting CP because of the positive relationship between content quality and congestion and the complementarity between content quality and network capacity.

The literature providing a formal economic analysis of zero-rating contracts is scant. To the best of our knowledge, no work has yet investigated its parallel implications for network investment and content quality provision with ambient congestion. Much of the literature models a monopolistic ISP connecting consumers to CPs, drawing exogenous revenues from advertising and competing passively to attract users. [Somogyi \(2017\)](#) considered the interplay between congestion and increasing utility from consumption under open and exclusive zero-rating contracts. In this model, the attractiveness of content plays a key role. When content is attractive, the ISP always offers an open zero-rating. [Jullien and Sand-Zantman \(2018\)](#) considered zero-rating contracts as an instrument to screen among traffic-sensitive CPs and to enhance allocative efficiency. In equilibrium, sponsored data are selected only by high-type CPs and improve network efficiency because they induce more traffic to be directed to high-valued content. Its welfare implications are ambiguous in that they depend on the mass of existing high-type CPs and the distribution of low-type CP values. Whereas in [Jullien and Sand-Zantman \(2018\)](#), the ISP uses zero-rating to screen among CPs, in [Inceoglu and Liu \(2019\)](#), zero-rating is implemented to screen among consumers in an environment with multiproduct demand. The ISP uses zero-rating to screen consumers according to the quantity consumed

---

<sup>9</sup>In [Choi and Kim \(2010\)](#), the rationale follows from the fact that CPs' willingness to pay for prioritization increases with congestion in the network. This feature is in line with empirical observations made by [Nevo et al. \(2016\)](#) and [Malone et al. \(2017\)](#).

and the composition of consumption. Zero-rating is found to be welfare-enhancing and cause network capacity expansion. [Jeitschko et al. \(2020\)](#) considered the implications of zero-rating with a vertically integrated ISP and asymmetric CPs with respect to some given quality parameter. [Schnurr and Wiewiorra \(2018\)](#), who analyzed two groups of consumers distinguished by their valuation of the content, find that when consumer groups are heterogeneous in their valuation of zero-rated content, they benefit from this practice. In contrast, when consumers are rather homogeneous in this preference, zero-rating might harm them. [Gautier and Somyi \(2020\)](#) compared the market outcomes under both zero-rating and paid prioritization with two CPs that are horizontally differentiated and asymmetric—the "stronger" CP has a larger natural market than its "weaker" counterpart. The general conclusion is that paid prioritization is preferable when traffic is valuable for CPs and congestion is severe—in the other cases, ISPs tend toward zero-rating. Finally, [Hoernig and Monteiro \(2020\)](#) studied the role of network effects in an ISP's rationale for implementing zero-rating. They noted that zero-rating is the profit-maximizing choice if network effects are strong enough and if the costs of increasing network capacity are low. They also noted that the result is similar under monopoly and duopoly conditions; however, in the latter case, the ISP with the larger consumer base benefits the most. We extend this literature by, first, finding that zero-rating makes content more expensive for consumers to access and implies a downward distortion of content quality by increasing downward vertical differentiation. Through zero-rating, the ISP is able to increase CPs' asymmetry and affect consumers' willingness to pay for content. CPs move from a minimal differentiation equilibrium to a downward vertical differentiation outcome. Second, we show that zero-rating is instituted to reduce congestion, whereas investments in broadband capacity are strictly lower than those under net neutrality. A complementarity effect between content quality and investment in network capacity is at play.

## 2 The model

We assume that a monopolistic ISP operates a broadband network through which CPs must deliver their services to end users. The ISP acts as a two-sided platform that connects CPs to a unit mass of consumers distributed uniformly on the line segment  $[0, 1]$ . Consumers are assumed to have unit demand. Two regulatory regimes are considered: a neutral network, which represents strict net-neutrality regulations (denoted  $n$  hereafter), and a discriminatory network, which represents zero-rating exclusive contracting (denoted  $z$  hereafter).

**Monopolistic ISP** Under both regimes, the ISP charges consumers a tariff  $\Phi = (H, \kappa, \tau_i)$ , which consists of a subscription fee (connection)  $H$ , a data allowance (data cap)  $\kappa$ , and an overage fee,  $\tau_i$ , charged in excess of  $\kappa$  for the consumption of content  $i = 1, 2$ . Then, to access content, a consumer with unit demand pays a total price of  $\mathcal{P} \equiv H + \tau \max\{0, 1 - \kappa\}$ . As in the remainder of this work, we only consider the effect of the overage fee  $\tau_i$ , and the data cap offered by the ISP is assumed to be set to  $\kappa = 0$ . Therefore, the tariff charged to consumers reduces to the two-part tariff  $\Phi = (H, \tau_i)$ . In a discriminatory network, the ISP can propose a zero-rating contract to a unique CP, which exempts users from the overage fee associated with the consumption of that content. We then have  $\tau_i = \tau_j \equiv \tau$  in a neutral regime, whereas we



have  $\tau_i = 0$  and  $\tau_j > 0$  under exclusive zero-rating when content  $i$  is zero-rated and content  $j$  is not.

The ISP provides a unique network capacity  $\mu_k$ , and the traffic generated on the network bears a cost for the ISP that is proportional to the network capacity provision. Let  $I(\mu_k)$  denote the cost of providing a given quality level  $\mu_k$ , with  $I' > 0$  and  $I'' > 0$ . Then, provided that both contracts are accepted, the ISP's payoff under regime  $k = n, z$  is given by

$$\Pi_k^{\text{ISP}} \equiv \begin{cases} H_n + \tau(D_1^n + D_2^n) - I(\mu_n) & \text{under uniform pricing,} \\ H_z + \tau_i D_i^z \mathbb{1}_{ZRj} - I(\mu_z) & \text{under zero-rating,} \end{cases}$$

with  $\mathbb{1}_{ZRj} = 1$  if CP  $j$  is zero-rated and 0 otherwise, where  $D_i^k$  represents the demand for content  $i$ . In a situation in which a CP is indifferent between its outside option (rejecting the zero-rating plan) and being part of the zero-rating program, we assume that it accepts the contract.

**Content providers** Two CPs,  $i = 1, 2$ , offer two distinct contents with quality  $q_i > 0$  and bear no cost associated with content production. Quality  $q_i$  interacts with the consumer's intrinsic valuation of content and affects his or her willingness to pay for content. Thus, the CPs are both horizontally and vertically differentiated. We consider an advertising-supported content model in which content is provided for free to consumers, and CPs compete on quality and advertising time to attract users. Consumers are subject to advertisements when they use content, and higher quality content results in higher ad exposure. This exposure to advertisements implies a disutility for consumers that is proportional to the time spent on the content, which is determined by its quality level. Let  $a_i$  be the advertising exposure level borne by a consumer. We say that visiting CP 2's content implies a higher exposure level if  $a_1 < a_2$  such that the CPs differ in their marginal advertising revenues. CPs' revenues from advertising are defined by a function  $R_i(a_i, s_i)$ , where  $s_i$  represents the price-per-click associated with each ad space sold by CP  $i$ . Let us make the following assumption:

**Assumption A1.** *Advertising revenues are proportional to the quality of the content offered, and the price-per-click is set equal to  $s_i = 1$ . As a result, CP  $i$ 's advertising revenues are  $R_i(a_i, s_i) = q_i a_i$ .*

Consequently, CP  $i$ 's payoff is given by

$$\pi_i^k(\cdot) = q_i^k a_i^k D_i^k(\cdot) - C(q_i), \quad (1)$$

where  $D_i^k(\cdot)$  denotes CP  $i$ 's market share under regime  $k = n, z$ . Finally, we make the following standard assumptions with respect to the investment in the quality cost function  $C(\cdot)$ :  $C'(\cdot) > 0$ ,  $C''(\cdot) > 0$  and  $C(0) = C'(0) = 0$ . Note that the total cost for a CP is assumed to be separable in quality (given by  $C$ ) and quantity (set equal to zero) such that the quality exhibits the characteristics of a public good for consumers.

**Consumers** The ISP charges consumer tariff  $\Phi = (H_n, \tau)$  in the neutral network and tariff  $\Phi = (H_z, \tau_i \mathbb{1}_{ZRj})$  in the discriminatory network. Each user has unit demand and single homes, and a consumer of type  $x \in [0, 1]$  that patronizes CP  $i$  has utility,

$$U_i \equiv \theta q_i^k - t|x - l_i| - \omega^k(\mathbf{x}, \mu_k) - q_i^k a_i^k - \mathcal{P}_k, \quad (2)$$

with

$$\mathcal{P}_k \equiv \begin{cases} H_n + \tau & \text{under uniform pricing,} \\ H_z + \tau_i \mathbb{1}_{ZRj} & \text{under zero-rating,} \end{cases}$$

where  $\theta$  denotes a gross surplus that is assumed to be large enough, and  $t$  is the transportation cost. Firm  $i$  is allocated position  $l_i$  on the unit line such that CP 1 is at  $l_1 = 0$  and CP 2 is at  $l_2 = 1$ . Then,  $t|x - l_i|$  denotes the disutility for consumers of type  $x$  from using content that is not their preferred horizontal specification. Finally,  $\omega(\mathbf{x}, \mu_k)$  denotes the level of congestion on the network supported by a consumer. As in [Calzada and Tselekounis \(2018\)](#), note that from equations (1) and (2), advertising exposure  $a_i^k$  acts as a unit price paid by consumers to CPs; hence, the results derived from our setup carry over to a subscription-based business model in which CPs charge a per-unit price for content.

Throughout this work, we make the following assumption, which ensures nonempty market share and CPs participation:

**Assumption A2.** *Consumer heterogeneity is such that  $9t > \theta^2 > 4t^2$ .*

The decision to patronize one CP over the other is constrained by the average level of congestion within the network  $\omega(\mathbf{x}, \mu_k)$ . Let  $\gamma_1 \in [0, 1]$  and  $\gamma_2 \in [0, 1]$  denote the request rate for content  $i = 1, 2$ , respectively, which are reflective of content attractiveness and capture the idea that consumers might ask each content several times in a given period. Parameter  $\gamma_i$  can be interpreted as the probability that content  $i$  is clicked on in a given period. To make zero-rating an attractive strategy for the ISP as a tool to discriminate among users, we make the following assumption.

**Assumption A3.** *The content provided by CP 1 (called the strong CP) is more attractive than the content provided by CP 2 (called the weak CP). Then, request rates are assumed to be such that  $\gamma_1 > \gamma_2$ .*

We then define the average congestion as follows:

$$\omega(n_1^e, n_2^e) \triangleq \frac{\gamma_1 n_1^e + \gamma_2 n_2^e}{\mu_k},$$

where  $(n_1^e, n_2^e)$  are the expected market shares of the CPs. The level of congestion is a decreasing function of the network capacity  $\mu$  and an increasing function of the content request rate  $\gamma_i$  and the total level of demand on the network. As we search for fulfilled expectations equilibria, we impose that  $n_1 = n_1^e = x$  and  $n_2 = n_2^e = (1 - x)$ , which is common in the literature and states that consumers rationally anticipate the decisions of other consumers such that the equilibrium location of the indifferent consumer is to be defined as a fixed point of the demand functions (e.g., [Economides and Tåg \(2012\)](#), [Choi and Kim \(2010\)](#), and [Gautier and Somogyi \(2020\)](#)).

Consumer surplus is given by the following quantity for  $i, j = 1, 2$ , and  $i \neq j$ ,

$$CS_k \triangleq \int_0^{D_1^k} U_1(z) dz + \int_{D_1^k}^1 U_2(z) dz, \quad (3)$$

with  $k = n, z$ . Social welfare is defined in the standard way as the gross benefits  $W_g^k$  from content net of the cost of the quality investment  $\sum_{i=1,2} C(q_i)$  and the cost of the network capacity

investment  $I(\mu)$ , with

$$W_g^k \triangleq \int_0^{D_1^{k*}} (\theta q_1^k - tz - \omega(z, \mu)) dz + \int_{D_1^{k*}}^1 (\theta q_2^k - t(1-z) - \omega(z, \mu)) dz. \quad (4)$$

**Notation** Throughout the text, we use the following notations:

$$\begin{aligned} \Delta_x &:= x_1 - x_2 & ; & & \Delta_{x_i} &:= x_i - x_j, \\ \Delta_{x_+} &:= x_1 + x_2 & ; & & \Delta_{x_{i+}} &:= x_i + x_j, \end{aligned}$$

for  $i = 1, 2$  and  $i \neq j$ .

**Timing** The timing of the game is as follows: at  $t = 1$ , the profit-maximizing ISP determines its investment level in congestion-reducing investments  $\mu_k$  at cost  $I(\mu_k)$ . At  $t = 2$ , CPs simultaneously determine their investment level in quality  $q_i^k$  at cost  $C(\cdot)$ .

At  $t = 3$ , under strict net neutrality, the ISP does not price discriminate between CPs and charges consumers the tariff  $(H_n, \tau)$  for each content. In the absence of net neutrality regulations, the ISP offers an exclusive zero-rating contract to CP  $i = 1, 2$ . If CP  $j$  accepts, the ISP charges consumers  $(H_z, \tau_i \mathbb{1}_{ZRj})$  and  $(H_z, 0)$  for non-zero-rated and zero-rated content, respectively. At  $t = 4$ , CPs either accept or reject the zero-rating contract and simultaneously choose the advertising exposure rate  $a_i^k$ . Finally, at  $t = 5$ , consumers either subscribe or not and decide on the content to consume.

The reason we assume that the ISP sets its network quality in stage 1 is that we assume that investments in the network are more of a long-run decision than the choice of content quality by CPs. The game is solved using backward induction, and we restrict the analysis to subgame perfect equilibria (SNPEs) in pure strategies.

### 3 Equilibrium analysis

In this section, we derive the subgame equilibrium of each market configuration and then compare the implications of a departure from net neutrality for market equilibrium values. We compare outcomes assuming that content and network qualities are exogenously given. In the neutral network, which corresponds to a strict net neutrality regulation, a profit-maximizing ISP can only use the network access fee to extract a surplus from the indifferent consumer. In the discriminatory network, which corresponds to an exclusive zero-rating contract, the ISP implements a per-unit fee in excess of the network fee to distort market competition and market demand.

#### 3.1 Market equilibrium

We begin this section by noticing that, given that we assume  $\kappa \equiv 0$ , uniform pricing is equivalent to setting the marginal fee  $\tau$  to zero.

**Lemma 1.** *Under net neutrality, charging consumers a nondiscriminatory tariff  $\Phi = (H_n, \tau)$  is equivalent to charging a single access price equal to  $H_n$ .*

Therefore, under a nondiscriminatory regime, the ISP charges a single price  $H_n$ , and under zero-rating, he charges the tariff  $\Phi = (H_z, \tau \mathbb{1}_{ZR_j})$  whenever content  $j$  is zero-rated. The following lemma states that the optimal strategy for the ISP at stage 2 is to enter into an agreement with the CP that enjoys the lowest content request rate  $\gamma_2$ , which we call the weak content provider.

**Lemma 2.** *Under the assumption that content request rates are such that  $\gamma_1 > \gamma_2$ , for the ISP to zero-rate the CP with the highest content request rate is not profitable<sup>10</sup>.*

Taking as given the investment decisions made by the ISP and CPs at the first and second stages of the game and assuming subgame perfection, we derive consumer demand; we then derive the optimal amount of advertisement to which CPs expose their users and, finally, the optimal pricing policy used by the ISP. The location  $\hat{x}_k$  of the marginal consumer who is indifferent between buying the content from either CP is the solution to

$$\theta q_1^k - t\hat{x}_k - a_1^k q_1^k - \tau \mathbb{1}_{ZR_2} = \theta q_2^k - t(1 - \hat{x}_k) - a_2^k q_2^k,$$

and is given by

$$\hat{x}_k = \frac{1}{2} + \frac{\theta(q_1^k - q_2^k) - a_1^k q_1^k + a_2^k q_2^k - \tau \mathbb{1}_{ZR_j}}{2t},$$

as consumers are uniformly distributed, the distribution of the market shares is given by  $D_1^k(\mathbf{a}^k, \mathbf{q}^k) \equiv \hat{x}_k$  and  $D_2^k(\mathbf{a}^k, \mathbf{q}^k) \equiv 1 - \hat{x}_k$ , respectively. Readily observed is that the ISP is able to distort market shares to the benefit of the contracting CPs by imposing a surcharge on users of the non-zero-rated content.

At the fourth stage of the game, the optimal pricing strategy for the ISP and the level of investments are taken into account by CPs when choosing the advertising exposure time that they impose on each of their users. CP  $i$ 's problem is then to find  $a_i^{k*}$  such that

$$a_i^{k*} \in \operatorname{argmax}_{a_i^k > 0} q_i^k a_i^k D_i^k(\mathbf{a}^k, \mathbf{q}^k) - C(q_i^k),$$

which yields the following as CP  $i$ 's best response:

$$\mathcal{BR}_1(a_2) = \frac{q_2^k(a_2^k - \theta) + \theta q_1^k + t - \tau \mathbb{1}_{ZR_j}}{2q_1^k}; \quad \mathcal{BR}_2(a_1) = \frac{q_1^k(a_1^k - \theta) + \theta q_2^k + t + \tau \mathbb{1}_{ZR_j}}{2q_2^k},$$

such that it is clear that advertising exposure rates are strategic complements. For given content qualities, advertising exposure rates (and ad nuisance) reflect a per-unit price charged to consumers. Therefore, CPs compete à la Bertrand in the amount of ads that they expose their users to, which is why best responses are upward sloping. Solving for both reaction functions yields the optimal amount,

$$a_1^{k*} = \frac{\theta(q_1^k - q_2^k) + 3t - \tau \mathbb{1}_{ZR_j}}{3q_1^k}; \quad a_2^{k*} = \frac{\theta(q_2^k - q_1^k) + 3t + \tau \mathbb{1}_{ZR_j}}{3q_2^k}. \quad (5)$$

<sup>10</sup>See the discussion following proposition 1. This follows from the restriction that we do not allow the ISP to charge a negative per-unit fee,  $\tau < 0$

**Neutral network** Given content qualities  $q_z^* = (q_1^{z*}, q_2^{z*})$ , the ISP chooses its tariff ( $H_n^*$ ) to extract the marginal consumer's surplus, that is,  $\hat{U}(\hat{x}_n, H_n^*) = 0$ , and sets the optimal connection fee to

$$H_n^* = \frac{\theta \Delta q_+^n}{2} - \frac{\Delta \gamma_+}{2\mu_n} - \frac{\theta \Delta \gamma \Delta q^n}{6t\mu_n} - \frac{3}{2}t.$$

Here,  $\Delta q_+^n = q_1^n + q_2^n$  and  $\Delta \gamma_+ = \gamma_1 + \gamma_2$ . Subgame equilibrium demand and advertisement levels are then given respectively by:

$$D_i^n(q_n) = \frac{1}{2} + \frac{\theta(q_i^n - q_j^n)}{6t} \quad \text{and} \quad a_i^* = \frac{2t}{q_i^n} D_i^n(q_n), \quad \text{for } i, j = 1, 2 \text{ and } i \neq j,$$

and profits are:

$$\pi_i^*(q_n) = \frac{(\theta(q_i^n - q_j^n) + 3t)^2}{18t} - C(q_i^n), \quad \text{for } i, j = 1, 2 \text{ and } i \neq j,$$

Under the constraint that both contents are offered at the same exogenous quality, set to  $q_1^n = q_2^n = q_n$ , CPs do not support the cost of producing content functionalities; that is,  $C(q_n) = 0$ , we obtain that  $D_i^n = \frac{1}{2}$ ,  $a_i^{n*} = \frac{t}{q_n}$ ,  $\pi_i^n = \frac{1}{2}t$ , and  $\Pi_n^{\text{ISP}} \equiv H_n^* = \theta q_n - \frac{3}{2}t$ . Then, under a neutral regime, CPs increase their advertising exposure as consumer heterogeneity increases, that is, when  $t$  increases, and reduce this amount through content differentiation. Higher transportation costs enhance each firm's market power, which allows for higher exposure.

**Discriminatory regime** For zero-rating, plugging  $a^{z*}$  into the demand and profit functions yields:

$$D_1^z = \frac{1}{2} + \frac{\theta(q_1^z - q_2^z)}{6t} - \frac{\tau}{6t} \quad \text{and} \quad D_2^z = \frac{1}{2} + \frac{\theta(q_2^z - q_1^z)}{6t} + \frac{\tau}{6t},$$

and

$$\pi_1^z = \frac{(\tau - 3t - \theta q_1^z + \theta q_2^z)^2}{18t} - C(q_1^z) \quad \text{and} \quad \pi_2^z = \frac{(\tau + 3t - \theta q_1^z + \theta q_2^z)^2}{18t} - C(q_2^z).$$

Assuming exogenous and symmetric content qualities, market demands reduce to  $D_1^z = \frac{1}{2} - \frac{\tau}{6t}$  and  $D_2^z = \frac{1}{2} + \frac{\tau}{6t}$ , advertisement reduces to  $a_1^{z*}(q_z, \tau) = \frac{t}{q_z} - \frac{\tau}{3t}$  and  $a_2^{z*}(q_z, \tau) = \frac{t}{q_z} + \frac{\tau}{3t}$ , and profit functions reduce to  $\pi_1^z = \frac{(\tau - 3t)^2}{18t}$ ,  $\pi_2^z = \frac{(3t + \tau)^2}{18t}$ . In contrast to the neutral regime, charging users of non-zero-rated content with a positive per-unit fee implies an asymmetric shift in the advertising exposure rate. The overage fee is partially absorbed by the non-zero-rated CP through  $a_1^{z*}(q_z, \tau)$ , which proportionally reduces the level at which its users are exposed to advertisement, whereas the zero-rated CP increases this amount accordingly because it benefits from the distortion in market shares. We then derive the following lemma:

**Lemma 3.** *Assuming symmetric content qualities, for a given per-unit fee, the more (less) content is vertically differentiated, the lower (higher) the ad exposure rate is for users.*

To observe the interaction between advertising exposure time and quality within the ISP's optimal pricing, let us assume that  $a^{z*} = (a_1^{z*}, a_2^{z*})$  and  $q^z = (q_1^z, q_2^z)$  are given. Again, the ISP

extracts all of the surplus from the marginal end user and sets  $H_z(\cdot)$  such that  $\hat{U}(\hat{x}_z, H_z(\cdot)) = 0$ , which leads to the following:

$$H_z^*(a^{z^*}, q^z, \tau) = \frac{\theta \Delta_{q^z}}{2} - \frac{a_1^{z^*} q_1^z + a_2^{z^*} q_2^z}{2} - \frac{\Delta_{\gamma^+}}{2\mu_z} - \frac{1}{2}t - \frac{\theta \Delta_{q^z} - a_1^{z^*} q_1^z + a_2^{z^*} q_2^z - \tau}{2\mu_z t} \Delta_{\gamma}, \quad (6)$$

where  $\Delta_{q^z} = q_1^z - q_2^z$ . Now, given  $H_z(q^z, a^{z^*}, \tau)$ , the ISP's problem is to find  $\tau^*$  such that

$$\tau^* \in \operatorname{argmax}_{\tau} H_z(q^z, a^{z^*}, \tau) + \tau D_i^z(q^z, a^{z^*}, \tau),$$

which, given  $a^{z^*}$  and  $q^z$ , leads to the following optimum overage fee<sup>11</sup>,

$$\tau^*(q^z, a^{z^*}, \mu) = \frac{\theta \Delta_{q^z}}{2} + \frac{a_2^{z^*} q_2^z - a_1^{z^*} q_1^z}{2} + \frac{\delta \Delta_{\gamma}}{2\mu_z}. \quad (7)$$

where  $\Delta_{\gamma} = \gamma_1 - \gamma_2$  and  $\Delta_{q^z} = q_1^z - q_2^z$ .

The optimum per-unit fee is increasing in the relative quality  $\Delta_{q^z}$  offered on the market and in the difference in advertising revenues drawn by the CPs. The ISP captures the residual rent left by the CPs through the difference in their advertising revenues through the overage fee. As a result, the greater the quality differentiation between contents and the greater the difference in ad exposure levels, the greater the per-unit fee charged by the ISP. An increase in content differentiation makes demand for content more inelastic, which implies that each CP has greater market power. This entails the possibility of charging more to consumers who purchase non-zero-rated content.

Now, plugging the expression for  $a_i^{z^*}$  given by (5) into (6) and (7), the ISP chooses its tariff  $(H_z^*, \tau^*)$  such that

$$\begin{aligned} (H_z^*, \tau^*) \in \operatorname{argmax}_{H_z, \tau > 0} H_z(q^z, a^{z^*}, \tau) + \tau D_i^z(q^z, a^{z^*}, \tau), \\ \text{s.t. } \hat{U}(\hat{x}, H_z^*, \tau) = 0, \end{aligned}$$

which yields

$$H_z^* = \frac{\theta \Delta_{q^z}}{2} - \frac{\Delta_{\gamma^+}}{2\mu_z} - \frac{\theta \Delta_{\gamma} \Delta_{q^z}}{6t\mu_z} - \frac{3}{2}t + \frac{\Delta_{\gamma}}{6t\mu_z} \tau + \frac{\tau}{2}, \quad (8)$$

$$\tau^*(q^z, \mu_z) = \frac{\theta \Delta_{q^z}}{2} + \frac{\delta \Delta_{\gamma}}{2\mu_z}, \quad (9)$$

where  $\Delta_{\gamma^+} = \gamma_1 + \gamma_2$ ,  $\Delta_{\gamma} = \gamma_1 - \gamma_2$ ,  $\Delta_{q^z} = q_1^z + q_2^z$  and  $\Delta_{q^z} = q_1^z - q_2^z$ .

Note that when no congestion exists and if firms are homogeneous by offering the same quality level, that is,  $q_i^z = q_j^z = q$ , then we recap the standard result that the optimal pricing policy for the ISP with homogeneous consumers has a per-unit fee equal to the marginal cost, which is assumed to be 0. Then, reminiscent of a price discrimination logic, greater content attractiveness or willingness to pay for content (through preference-matching  $\theta q_i^z$ ), greater difference in content request  $\Delta_{\gamma}$ , and greater degree of vertical differentiation  $\Delta_{q^z}$  translate into a higher per-unit fee on non-zero-rated content. Note also that the optimum overage fee is convex in network capacity. This highlights the fact that a profit-maximizing ISP can strategically reduce its investments in broadband capacity to extract more rent from

<sup>11</sup>A sufficient condition for the concavity of the profit function is  $\theta(q_1^z - q_2^z) > 2\tau - \frac{\delta \Delta_{\gamma}}{\mu_z}$ .

the consumer side because scarcer (more abundant) resources result in higher (lower) per-unit surcharge. Then, the optimal pricing policy from the ISP leads the marginal user to be located at

$$\hat{x}_z^* = \frac{1}{2} + \frac{\theta \Delta_{q_i^z}}{12t} - \frac{\Delta_{\gamma}}{12t\mu_z}, \quad (10)$$

which, assuming exogenous qualities, yields  $\hat{x}_z^* = \frac{1}{2} - \frac{\Delta_{\gamma_i}}{12t\mu_z}$  and to advertising exposure rates

$$a_i^{n*} = \frac{t}{q_i^n} + \frac{\theta \Delta_{q_i^n}}{3q_i^n} \quad \text{and} \quad a_i^{z*} = \frac{t}{q_i^z} + \frac{\theta \Delta_{q_i^z}}{6q_i^z} - \frac{\Delta_{\gamma_i}}{6\mu_z q_i^z},$$

where  $\Delta_{q_i^k} = q_i^k - q_j^k$  and  $\Delta_{\gamma_i} = \gamma_i - \gamma_j$  for  $i, j = 1, 2, i \neq j$  and  $k = n, z$ .

Whereas CPs share the market under net neutrality, financially differentiated content clearly enables the profit-maximizing ISP to distort market competition through the difference in content requests, and this distortion is more likely to be high for lower network capacity.

### 3.2 Effect on advertising exposure rates, access prices, and profits

In this section, we compare the effect of a departure from net neutrality to zero-rating under the assumption that content quality is exogenously given, that is,  $q_i^k = q_k$  with  $k = n, z$ . We begin by stating the main proposition that we derive from this subsection and then consider each case in the following.

**Proposition 1.** *Suppose that content quality is exogenous, that is,  $q_i^k = q_k$  with  $k = n, z$ , and that a profit-maximizing ISP can financially discriminate between contents. Then,*

- (i) *the zero-rated CP exposes its consumers to advertisements more than its non-zero-rated competitor;*
- (ii) *the weak CP is zero-rated, the ISP internalizes the disutility from congestion in its access price, and the subscription fee is greater under net neutrality than under zero-rating agreements;*
- (iii) *the greater the vertical differentiation and the difference in ad exposure, the higher the per-unit fee is; and*
- (iv) *a departure from net neutrality increases the ISP's profits and content industry's surplus and strictly reduces the noncontracting CP's profits.*

**Effect on advertising exposure.** Part (i) of the proposition states that the optimum ad exposure is asymmetric between CPs and that the zero-rated content is subject to a greater degree of advertisement. Indeed, assuming that content quality is exogenously given, the optimum amount of advertising exposure in both regimes is given by

$$a_i^{n*}(q_n) = \frac{t}{q_n}, \quad (11)$$

and

$$a_i^{z*}(q_z, \mu_z) = \frac{t}{q_z} - \frac{\gamma_i - \gamma_j}{6q_z\mu_z}. \quad (12)$$

The optimum ad exposure depends on the magnitude of the difference in content request rates. Specifically, the advertising time embodied in the non-zero-rated content is decreasing in the difference in content request rates, whereas the zero-rated CP increases the time exposure because this difference increases. In a neutral network, CPs share the market, that is,  $x_i^n = 1/2$ , and expose consumers to the same amount of advertising. However, because the surcharge charged by the ISP acts as a standard excise tax, observing that the non-zero-rated CP supports a reduction in its demand is straightforward, whereas the zero-rated content supports an increase in its own, that is,  $x_1^z < x_1^n = x_2^n < x_2^z$ . This finding implies a shift in the advertising exposure rates of both CPs, and we can verify that users of the contracting CP are more exposed than users of the strong CP in a discriminatory network, that is,  $a_2^z > a_1^z$ . This makes sense because the zero-rated CP benefits from an increase in its demand through the zero-rating contract, which allows it to charge consumers more. In response, its non-zero-rated competitor is forced to reduce its per-unit advertisement rate, which is why we observe this relation between prices under the two regimes.

**Effect on the ISP's pricing policy.** A general claim against net neutrality is that because it forbids charging CPs to access the network, a profit-maximizing ISP cannot extract revenues from CPs, which could ultimately lead to a higher network access fee for consumers. In line with this claim, part (ii) of the proposition states that, with respect to uniform pricing in the neutral network, content is cheaper to access when the ISP is allowed to impose a positive per-unit fee than content under net neutrality because the ISP reduces it to accommodate both the congestion effect and total advertising exposure.

To see why, consider first the ISP's pricing policies under the two regimes when advertisement levels are taken as given. Under net neutrality, we obtain the following subscription fee:

$$H_n(a^n, q_n, \mu_n) = \theta q_n - \frac{t}{2} - \frac{\Delta_{a_+^n}}{2} q_n + \frac{\Delta_{a^n} \Delta_\gamma}{2t\mu_n} q_n - \frac{\Delta_{\gamma_+}}{2\mu_n},$$

and under zero-rating agreements, the subscription and the per-unit fees are given by

$$H_z(a^z, q_z, \mu_z) = \theta q_z - \frac{t}{2} - \frac{\Delta_{a_+^z}}{4} q_z + \frac{\Delta_{a^z} \Delta_\gamma}{4t\mu_z} q_z - \frac{\Delta_{\gamma_+}}{4\mu_z} - \frac{a_2^z}{2} q_z - \frac{\gamma_1}{2\mu_z},$$

and

$$\tau = \frac{\Delta_\gamma}{2\mu_z} - \frac{\Delta_{a^z}}{2} q_z,$$

with  $\Delta_\gamma = \gamma_1 - \gamma_2$ ,  $\Delta_{a_+^k} = a_1^k + a_2^k$  and  $\Delta_{a^k} = a_1^k - a_2^k$ .

Thus, the network access fee is decreasing in the total ad exposure level and is increasing in content vertical differentiation and quality complementarity. The greater the vertical differentiation and difference in ad exposure, the higher the per-unit fee, which is part (iii) of the proposition. The fact that the connection fee decreases as the total amount of advertising exposure increases also holds when the ISP is allowed to contract with the weak CP. However, by contracting with the weak CP, the ISP is able to extract the residual rent,  $(a_2^z - a_1^z)$ , left by the CPs through the optimum per-unit surcharge imposed on consumers. Overall, under both regulatory regimes, the ISP reduces the network access fee to accommodate the total advertising time to which consumers are exposed.



Now, if we introduce the level of advertising exposure derived previously and plug expressions (11) and (12) into  $H_n(a^n, q_n, \mu_n)$  and  $H_z(a^z, q_z, \mu_z)$ , we obtain

$$H_n(q_n, \mu_n) = \theta q_n - \frac{3}{2}t - \frac{\Delta\gamma_+}{2\mu_n},$$

$$H_z(q_z, \mu_z) = \theta q_z - \frac{3}{2}t - \frac{\Delta\gamma_+}{4\mu_z} + \frac{\Delta\gamma^2}{12t\mu_z} - \frac{1}{2\mu_z}\gamma_1,$$

which can be written more intuitively by using the expressions for the subgame equilibrium levels of congestion in the network.

Indeed, assuming CPs that are symmetric in quality and taking ad exposure levels as given, the average levels of congestion under the two regimes are given by

$$\omega_n(a^n, q_n, \mu_n) = \frac{\Delta\gamma_+}{2\mu_n} - \frac{\Delta a^n \Delta\gamma}{2t\mu_n} q_n, \quad (13)$$

$$\omega_z(a^z, q_z, \mu_z) = \frac{\Delta\gamma_+}{2\mu_z} - \frac{\Delta a^z \Delta\gamma}{2t\mu_z} q_z - \frac{\Delta\gamma^2}{4t\mu_z^2}, \quad (14)$$

from which it is observed that the average level of congestion strictly increases with the total number of content requests, which is quite intuitive, and decreases in the difference in the request and ad exposure levels under the two regimes. Using expressions (11) and (12), the subgame equilibrium levels of congestion are then given by

$$\omega_n(\mu_n) = \frac{\Delta\gamma_+}{2\mu_n}, \quad (15)$$

$$\omega_z(\mu_z) = \frac{\Delta\gamma_+}{2\mu_z} - \frac{\Delta\gamma^2}{12t\mu_z^2}. \quad (16)$$

Note that whenever a profit-maximizing ISP is allowed to financially discriminate between the two contents and content quality is symmetric, the average congestion level is lower under a discriminatory rather than a neutral network, which is consistent with the literature on net neutrality (e.g., Bourreau et al. (2015)). Then, plugging the expressions given by (15) and (16) into  $H_n(q_n, \mu_n)$  and  $H_z(q_z, \mu_z)$ , we can write the subscription fees as

$$H_n(q_n, \mu_n) = \theta q_n - \frac{3}{2}t - \omega_n(\mu_n),$$

$$H_z(q_z, \mu_z) = \theta q_z - \frac{3}{2}t - \omega_z(\mu_z) - \frac{\Delta\gamma}{4\mu_z},$$

Thus, the ISP internalizes the disutility from congestion incurred by consumers on its network. Note that, assuming given qualities and congestion, the ISP can increase the access price in the discriminatory network by reducing network capacity<sup>12</sup>. Computing the difference  $H_n(q_n, \mu_n) - H_z(q_z, \mu_z)$ , we obtain that  $H_n(q_n, \mu_n) > H_z(q_z, \mu_z)$  if,

$$0 < \omega_n(\mu_n) - \omega_z(\mu_z) < \theta(q_n - q_z) + \frac{\Delta\gamma}{4\mu_z},$$

meaning that if quality levels are exogenously provided in both regimes, the assertion in part (ii) follows.

<sup>12</sup>This rent-extraction effect is at play in driving the ISP's incentives to invest in capacity, as is observed in the next section.

Finally, assuming given content qualities and zero-rating, for contracting with the strong CP (with the highest content request rate) to be profitable for the ISP, the optimal per-unit fee has to be negative. Denoting by  $\Pi_{ZR1}^{ISP}$ , the ISP profits when the strong CP is zero-rated and by  $\Pi_{ZR2}^{ISP}$  when the weak CP is zero-rated, the difference in the ISP profits is given by

$$\Pi_{ZR1}^{ISP} - \Pi_{ZR2}^{ISP} = -\frac{1}{3\mu_z t} \tau (\gamma_1 - \gamma_2)$$

which is positive iff  $\tau < 0$ .

**Effect on the ISP, content industry, and CPs' profits.** Let us now derive how the equilibrium profits of the ISP, CPs, and the content sector are affected by a move to zero-rating. The corresponding ISP equilibrium profits under the two regimes are given by

$$\begin{aligned} \Pi_n^{ISP} &= \theta q_n - \frac{3}{2}t - \omega_n(\mu_n) - I(\mu_n), \\ \Pi_z^{ISP} &= \theta q_z - \frac{3}{2}t - \omega_z(\mu_z) - \frac{\Delta_\gamma^2}{24t\mu_z^2} - I(\mu_z), \end{aligned}$$

which, using relations (15) and (16), writes as,

$$\begin{aligned} \Pi_n^{ISP} &= \theta q_n - \frac{3}{2}t - \frac{\Delta_{\gamma+}}{2\mu_n} - I(\mu_n), \\ \Pi_z^{ISP} &= \theta q_z - \frac{3}{2}t - \frac{\Delta_{\gamma+}}{2\mu_z} + \frac{\Delta_\gamma^2}{24t\mu_z^2} - I(\mu_z), \end{aligned}$$

From this, we can observe that the ISP always benefits from a discriminatory network for given qualities and network capacities. With respect to CPs' profits, in a neutral network, we obtain the following expressions for exogenous quality levels:

$$\pi_i^n = \frac{1}{2}t,$$

and

$$\pi_i^z = \frac{1}{2}t - \frac{(12t\mu_z - \Delta_{\gamma_i})}{72t\mu_z^2} \Delta_{\gamma_i}, \quad i = 1, 2,$$

under zero-rating, where  $\Delta_{\gamma_i} = \gamma_i - \gamma_j$  for  $i = 1, 2$  and  $i \neq j$ . The contracting CP benefits from a zero-rating agreement and realizes higher profits than its competitor, while for  $\mu_z > \tilde{\mu}_z = \frac{(\gamma_1 - \gamma_2)}{12t}$ , zero-rating strictly reduces the profits of the noncontracting CP. The content sector's overall profits, given by  $\Pi_k^{CPs} = \sum_i \pi_i^k$ , are

$$\Pi_n^{CPs} = t \quad \text{and} \quad \Pi_z^{CPs} = t + \frac{(\gamma_1 - \gamma_2)^2}{36t\mu_z^2},$$

and the assertion in part (iv) of proposition 1 follows.

## 4 Investments in content quality and network capacity

In this section, we consider the CPs' optimal decision to invest in their content attributes in stage 2 of the game and show that a discriminatory regime softens quality competition. We

then consider the effect that a marginal increase in content functionalities has on the degree of advertisement to which CPs expose their users and derive the subgame equilibrium level of advertisement in each regime given optimal content qualities. We find that advertising levels are symmetric between CPs and greater when zero-rating is implemented. Finally, we conclude this section by considering the optimal level of network capacity investments by the ISP. We find that when moving to zero-rating, it is optimal for the ISP to reduce network capacity because the existing strategic complementarity between content quality and network capacity reduces its network management costs in equilibrium.

#### 4.1 Content quality

Given the ISP's optimal pricing policy, the CPs simultaneously choose the level of quality of their content. We find that a departure from net neutrality implies an asymmetric equilibrium in content quality and that CPs have lower incentives to invest under zero-rating agreements.

The incentives to invest in content quality are then driven by a change in profits due to a marginal increase in the quality level. The first-order condition of CP  $i$ 's problem under regime  $k = n, z$  is given by

$$q_i^k \left[ \frac{\partial a_i^{k*}(\cdot)}{\partial q_i^k} D_i^{k*}(\cdot) + \frac{\partial D_i^{k*}(\cdot)}{\partial q_i^k} a_i^{k*}(\cdot) \right] + a_i^{k*} D_i^{k*} - C'(q_i^k) = 0. \quad (17)$$

For each regime, we then obtain the following marginal effects of a quality improvement on market demand:

$$\frac{\partial D_i^n}{\partial q_i^n} = \frac{\theta}{6t} \quad \text{and} \quad \frac{\partial D_i^z}{\partial q_i^z} = \frac{\theta}{12t'}, \quad (18)$$

and on the degree of advertising exposure

$$\frac{\partial a_i^n}{\partial q_i^n} = \frac{\theta q_i^n - 3t}{3q_j^{n^2}} \quad \text{and} \quad \frac{\partial a_i^z}{\partial q_i^z} = \frac{\mu_z (\theta q_j^z - 6t) + \gamma_i - \gamma_j}{6\mu_z q_i^{z^2}}, \quad (19)$$

for  $i = 1, 2$  and  $i \neq j$ . CPs are always able to increase their outputs by investing in the quality of their content, and the associated marginal benefits are driven by positive market share responses and, under both regimes given the sufficient condition that  $q_i^k < \frac{3t}{\theta}$ , by a decrease in advertising exposure levels.

We now compute the optimum levels of content quality for both regulatory regimes. We first derive CP  $i$ 's best response,  $\mathcal{BR}_i^n(q_j^n)$ , to a quality improvement from CP  $j$ ; then, we compute the equilibrium qualities and finally compare quality provision between CPs and across regimes. The first-order condition in (17) yields the best response for CP  $i$  in each regulatory regime,

$$\mathcal{BR}_i^n(q_j^n) = \frac{\theta(\theta q_j^n - 3t)}{\theta^2 - 9t} \quad \text{and} \quad \mathcal{BR}_i^z(q_j^z) = \frac{\theta\mu_z(\theta q_j^z - 6t) + \theta\Delta\gamma_i}{\mu_z(\theta^2 - 36t')},$$

for  $i = 1, 2$  and  $i \neq j$  and in which again  $\Delta\gamma_i = \gamma_i - \gamma_j$ . The nature of the strategic interaction between the two CPs is then given by the signs of

$$\frac{\partial \mathcal{BR}_i^n(q_j^n)}{\partial q_j^n} = -\frac{\theta^2}{9t - \theta^2} \quad \text{and} \quad \frac{\partial \mathcal{BR}_i^z(q_j^z)}{\partial q_j^z} = -\frac{\theta^2}{36t' - \theta^2},$$

making clear that CPs' qualities are strategic substitutes, as best responses are downward-sloping. Let  $p_i = a_i q_i$  be the total per-unit price of content  $i = 1, 2$ . Anticipating symmetric competition in advertising exposure rates, CPs use investment in content quality to influence the degree of "price" competition and to distort the asymmetry between them. Specifically, under zero-rating, the weak content provider (which is zero-rated) anticipates aggressive behavior in response to the non-zero-rated content provider. Therefore, the weak CP increases its quality to deter the aggressive investment of the strong CP, which implies (as shown below) that users of the zero-rated content are charged a higher total price.

Solving for both best responses and comparing incentives to invest in functionalities under each regime, we observe the following:

**Lemma 4.** *Suppose that  $C(q_i^k) = \frac{1}{2}q_i^{k^2}$ . Then, for  $i, j = 1, 2$  and  $i \neq j$ , the optimum quality in a neutral network is given by*

$$q_i^{n*} = q_n^* = \frac{\theta}{3},$$

and in a discriminatory network it is given by

$$q_i^{z*} = \frac{1}{2}q_n^* - \mathcal{A}_i(\theta, t, \mu_z),$$

where  $\mathcal{A}_i(\theta, t, \mu_z) \equiv \frac{1}{2} \left( \frac{\theta \Delta_{\gamma_i}}{\mu_z(18t - \theta^2)} \right)$  and  $\Delta_{\gamma_i} = \gamma_i - \gamma_j$ .

With this lemma, investments in content functionalities are shown to be misaligned between CPs when moving from a neutral to a discriminatory network. In the neutral network, CPs share the market in symmetric equilibrium, and the incentives to invest in quality are aligned between them because CPs invest the same amount in their content attributes/quality. In equilibrium, the content quality increases in consumers' intrinsic utility  $\theta$  but is independent of the network quality provided by the ISP. Zero-rating agreements distort CPs' market shares, giving them the opportunity to influence demand through the quality of their content. In a discriminatory regime, content quality is lower and asymmetric, but the effect of increasing content features is ambiguous. Quality competition depends on the ISP's pricing strategy and its investments in network capacity, and the expressions for  $q_i^{z*}$  reflect the amplifying effect that  $\mu_z$  can have on vertical content differentiation. This level still increases in consumers' intrinsic utility  $\theta$ , but this time, it decreases when network capacity grows. As a result, a marginal network capacity expansion increases the content quality of the non-zero-rated CP and reduces the zero-rated CP's content quality.

We now state the main result of this section. First, note that given assumption **A2**, we have that  $\mathcal{A}_1(\theta, t, \mu_z) > 0$  and  $\mathcal{A}_2(\theta, t, \mu_z) < 0$ , such that we immediately obtain  $q_1^{z*} < q_2^{z*}$ . Next, we observe that  $q_n^* = q_1^{z*} + q_2^{z*}$ , which implies that for  $i = 1, 2$ ,  $q_i^{z*} < q_n^*$ . As a result, these two observations allow us to derive the following immediate result:

**Proposition 2.** *Suppose that a profit-maximizing platform can financially discriminate between contents. Then, the non-zero-rated CP invests less in content quality than its zero-rated competitor, and both CPs underinvest with respect to a neutral network.*

The main message from this proposition is that zero-rating decreases the overall supply of quality content in the market. The more attractive content is, that is, the greater the functionalities that it has, the more congestion is supported at a given level of quality, and the more

consumers are exposed to advertising. By improving its content functionalities, the zero-rated CP induces a positive market share response only if this improvement reduces the disutility in congestion supported by its users, which is the case in equilibrium.

However, by opting for quality degradation, the noncontracting CP reduces the ISP's ability to price discriminate between users. Indeed, the noncontracting CP makes its content more attractive by increasing its quality, which implies a higher overage fee charged to its users by the ISP. Moreover, the zero-rated CP also increases the attractiveness of its content by improving its content quality, which implies that more users are willing to switch to it. Such a switch implies a strict loss for the ISP, which then—all things being equal—reduces the overage fee charged to these users to minimize the switching incentives. Hence, investments in quality by weak CP imply a decrease in the overage fee. In return, by degrading its own quality, the strong CP is able to decrease the per unit overage fee and the disutility from the congestion of its users because it actually reduces the congestion level by doing so—but does not sufficiently compensate for the switch. Then, the marginal effect on the profits of an increase in content quality is negative for the noncontracting CP, and the CP opts for quality degradation because such an investment is insufficient to eliminate users' switching incentives implied by the zero-rating agreement.

As a result and in contradiction to the main argument of opponents of net neutrality, the policy implication of the last proposition is that allowing for discriminatory practices is not a condition to foster content innovation by CPs. Because the investment level implied by the zero-rated CP does not sufficiently compensate for the lack of investment by its rival, overall, the content sector is less innovative. CPs switch from a minimal vertical differentiation outcome under net neutrality to an asymmetric equilibrium with a greater degree of downward vertical differentiation when zero-rating is allowed.

## 4.2 Optimal advertisement level

Recall that from relation (19), under the condition that  $q_i^k < \frac{3t}{\theta}$ , the marginal benefits from a content quality improvement are driven by a decrease in advertising exposure by CPs in that exposure rates are negative functions of content quality. Hence, we can state the following:

**Lemma 5.** *There exists a network capacity threshold  $\bar{\mu}_z(\theta, t)$  such that, following a marginal increase in content quality, the contracting CP always reduces the advertising exposure level embodied in its content, whereas the noncontracting CP reduces this level only if  $\mu_z > \bar{\mu}_z(\theta, t)$ .*

Let us define the network capacity threshold  $\bar{\mu}_z(\theta, t) \equiv \frac{3\Delta_\gamma}{18t - \theta^2}$ , with  $\Delta_\gamma = \gamma_1 - \gamma_2$ , and show that condition  $q_i^k < \frac{3t}{\theta}$  is indeed satisfied in both regulatory regimes. From relation (19) and lemma 4, it is straightforward that  $q_n^* < \frac{3t}{\theta}$ . Indeed, given  $q_n^*$ , this condition implies that  $\theta^2 < 9t$ , which is satisfied by assumption A2; hence,  $\frac{\partial a_i^k}{\partial q_i^k} < 0$ . Under a discriminatory regime, we obtain that  $\frac{\partial a_1^z}{\partial q_1^z} > 0$  if  $q_2^z > \frac{6t}{\theta} - \frac{\Delta_\gamma}{\theta\mu_z}$  and  $\frac{\partial a_2^z}{\partial q_2^z} > 0$  if  $q_1^z > \frac{6t}{\theta}$ . From proposition 2, we have that  $q_i^{z*} < q_n^* < \frac{3t}{\theta}$ ; hence, the condition for  $q_1^z$  cannot be satisfied, and we necessarily have that  $\frac{\partial a_2^z}{\partial q_2^z} < 0$ . Next, we have that  $q_2^z < \frac{\theta}{3} < \frac{6t}{\theta} - \frac{\Delta_\gamma}{\theta\mu_z}$  if  $\mu_z > \bar{\mu}_z(\theta, t)$ . Hence, for a sufficient degree of network capacity,  $\frac{\partial a_1^z}{\partial q_1^z} < 0$ , and the result in the lemma follows.

As a result, given both proposition 2 and lemma 5, allowing a profit-maximizing ISP to set a per-unit surcharge on users of higher quality content implies greater advertising exposure

in both market segments. We obtain the following result:

**Proposition 3.** *A departure from net neutrality increases the advertising exposure of both contents. In particular, the optimum advertising exposure rates is given by*

$$a_i^{n*} = \frac{3t}{\theta},$$

and

$$a_i^{z*} = \frac{6t}{\theta}.$$

The main message from this proposition is that zero-rating relaxes the competition for advertising nuisance and increases the degree of advertising to which users are exposed and the price per unit of the quality that users implicitly pay to access each content.

Under network neutrality, the ISP does not affect competition between CPs, whereas advertising competition is relaxed in the discriminatory equilibrium, which explains why users are more exposed to ads. In both regimes, advertising exposure rates are symmetric between market segments and independent of network capacity. Symmetry arises from the fact that CPs compete with each other à la Bertrand for "advertising nuisance" because the best responses are upward sloping, and independence from network capacity arises from quality competition between CPs. Indeed, an ISP that offers a zero-rating contract to the weak content provider reduces the asymmetry between content providers, which intensifies the competition in quality, through which content providers anticipate the ISP's investment strategy. ISP-induced distortion is counteracted by CPs' quality investments, further relaxing advertising competition. Therefore, in equilibrium, the effect of network capacity is already captured and mitigated upstream in the investment in content quality, making the marginal revenues associated with advertising exposure independent of network capacity<sup>13</sup>. Consumers support greater advertisement than in a neutral network, and market shares are given by

$$D_i^{z*}(\theta, t, \mu_z) = \frac{1}{2} - \frac{3(\gamma_i - \gamma_j)}{2\mu_z(18t - \theta^2)}, \text{ for } i, j = 1, 2 \text{ and } i \neq j.$$

Finally, note that the per-unit advertising exposure rate in our specification can be interpreted as a direct payment from consumers to CPs. Therefore, the model also exhibits the same properties as purely paid content, which consumers pay to access. The result in the last proposition highlights the fact that zero-rating reduces advertising competition between CPs but makes the zero-rated content more expensive for consumers to access. More precisely, letting  $p_i^k \equiv a_i^k q_i^k$  for  $k = n, z$ , we can derive the following corollary:

---

<sup>13</sup>The simple expressions for the optimal level of advertising and content quality obtained previously are because quality linearly affects the marginal willingness to pay of end-users without influencing users' transport costs and because users are homogeneous. Although this involves complex indirect effects, the same qualitative results can be shown to hold in a more general model in which, for example, the interaction between vertical and horizontal differentiation is such that quality interacts with distance-based utility (see [Anderson et al., 2017](#)). Here, increasing certain content attributes is not valued equally by users, and a shorter distance between content and user preference results in the content being more valued. In such a context, the quality of the content affects users' willingness to pay in a nonlinear way and, as long as users are of the same types and advertising enters the utility function as a price per unit of quality provision, the same qualitative results hold.

**Corollary 1.** *A departure from net neutrality increases the per-unit access price of weak content and reduces the per-unit access price of strong content compared with net neutrality. In particular,*

$$p_1^{z*} = t - \frac{3t(\gamma_1 - \gamma_2)}{(18t - \theta^2)\mu_z} < p_1^{n*} = t,$$

and

$$p_2^{z*} = t + \frac{3t(\gamma_1 - \gamma_2)}{(18t - \theta^2)\mu_z} > p_2^{n*} = t.$$

### 4.3 Investment incentives of the ISP

In this section, we consider the ISP's decision to invest in network capacity under both regulatory regimes. To assess whether a discriminatory regime generates greater incentives than a neutral network to invest in network capacity, we first consider the effect of a marginal increase in network capacity on equilibrium access prices, market shares, content qualities, and profits. We then derive the equilibrium network capacity in the neutral network, compare incentives to invest under zero-rating agreements, and find that a profit-maximizing ISP strategically reduces its network capacity through two channels.

Recall that for given content qualities  $q$  and network capacity  $\mu$ , the ISP's profits under net neutrality are

$$\Pi_n^{\text{ISP}}(q, \mu) = \frac{\theta q_1 + \theta q_2}{2} - \frac{\theta(q_1 - q_2)}{6t\mu} \Delta_\gamma - \frac{\Delta_{\gamma^+}}{2\mu} - \frac{3}{2}t - I(\mu), \quad (20)$$

and under zero-rating agreements, they are

$$\Pi_z^{\text{ISP}}(q, \mu) = \Pi_n^{\text{ISP}}(q, \mu) + \mathcal{B}(\theta, q, \mu), \quad (21)$$

where

$$\mathcal{B}(\theta, q, \mu) \equiv \frac{2\theta^2 \Delta_q^2 \mu^2 + \theta \Delta_q \Delta_\gamma \mu + \Delta_{\gamma^+}^2}{24t\mu^2}.$$

**Lemma 6.** *A marginal network investment (i) increases the access price in a neutral regime and increases the access price in a discriminatory network for all  $\mu > \mu(\theta, t)$ , (ii) does not affect market shares in a neutral regime but increases the market share of the noncontracting CP, and (iii) increases the content quality of the noncontracting CP and decreases the contracting CP's content quality.*

*Proof.* For part (i), the respective derivatives of the network access fee with respect to the network capacity in each regime are given by:

$$\frac{\partial H_n}{\partial \mu_n} = \frac{\gamma_1 + \gamma_2}{2\mu_n^2} > 0,$$

$$\frac{\partial H_z}{\partial \mu_z} = \frac{\mu_z (\gamma_1 (27t - 2\theta^2) + 9\gamma_2 t) - 6(\gamma_1 - \gamma_2)^2}{2(18t - \theta^2)\mu_z^3}.$$

Let

$$\mu(\theta, t) = \frac{6(\gamma_1 - \gamma_2)^2}{\gamma_1(27t - 2\theta^2) + 9\gamma_2 t},$$

then,  $\frac{\partial H_z}{\partial \mu_z} > 0$  if  $\mu_z > \mu(\theta, t)$ . For part (ii), the derivative of the noncontracting CP's market share with respect to the network capacity under zero-rating is given by:

$$\frac{\partial D_1^z}{\partial \mu_z} = \frac{3(\gamma_1 - \gamma_2)}{2(18t - \theta^2)\mu_z^2} > 0,$$

and for part (iii), the respective derivatives of the CPs' profits with respect to the network capacity in a discriminatory regime are given by:

$$\begin{aligned} \frac{\partial q_1^z}{\partial \mu_z} &= \frac{(\gamma_1 - \gamma_2)\theta}{2(18t - \theta^2)\mu_z^2} > 0, \\ \frac{\partial q_2^z}{\partial \mu_z} &= -\frac{(\gamma_1 - \gamma_2)\theta}{2(18t - \theta^2)\mu_z^2} < 0, \end{aligned}$$

and the assertions in the lemma follow.  $\square$

We now derive the optimal level of network capacity under each regime and then compare their levels with the socially optimal capacity. For this purpose, let us assume that the ISP is subject to the following cost of providing a given network capacity,  $I(\mu) \equiv \frac{k}{2}\mu^2$ , with  $k > 0$ . Taking into account the optimal content quality investment by CPs, the ISP's profits under each regime for a given network capacity are given by

$$\Pi_n^{\text{ISP}} = \frac{1}{3}\theta^2 - \frac{3}{2}t - \frac{\Delta\gamma_+}{2\mu} - I(\mu), \quad (22)$$

and

$$\Pi_z^{\text{ISP}} = \Pi_n^{\text{ISP}} + \frac{1}{6}[\mathcal{S}(\theta, t, \mu) - \theta^2], \quad (23)$$

where

$$\mathcal{S}(\theta, t, \mu) \equiv \frac{81t\Delta\gamma^2}{(\theta^2 - 18t)^2\mu^2} > 0.$$

For the neutral network, solving for the first-order condition of the expression for  $\Pi_n^{\text{ISP}}$ , we obtain that the optimum network capacity in a neutral network is given by

$$\mu_n^* = \frac{\Delta\gamma_+^{1/3}}{2^{1/3}k^{1/3}}.$$

We state the main result of this section:

**Proposition 4.** *A profit-maximizing ISP invests less in broadband infrastructure under zero-rating agreements than in a neutral network. The incentives depend on a rent-extraction and a cost-alleviation effect.*

The proof is straightforward, as from (22) and (23), we obtain that  $\left. \frac{d\Pi_n^{\text{ISP}}}{d\mu_n} \right|_{\mu_n=\mu} - \left. \frac{d\Pi_z^{\text{ISP}}}{d\mu_z} \right|_{\mu_z=\mu} = -\left. \frac{d\mathcal{S}(\theta, t, \mu)}{d\mu_z} \right|_{\mu_z=\mu} > 0$ . Although investments in capacity increase the demand for non-zero-rated content, they actually decrease the additional revenue that the ISP can extract from these consumers through the per-unit fee. The profit-maximizing ISP's decision to invest in broadband capacity is inversely related to the distortion in market share that doing so creates. Then, the more the ISP distorts demand to the benefit of the zero-rated content, the less incentive it has to improve the network and reduce congestion.



The rent-extraction effect is not new and relies on the attractiveness of the content. In equilibrium, even if the overall quality is lower in the content market, the contracting CP is vertically more differentiated than its competitor, making its content more attractive. Hence, consumers spending more time on this content reduces the capacity of the ISP to extract rent via the per-unit fee  $\tau^*(\cdot)$  charged to consumers using non-zero-rated content. To avoid losses, the ISP has incentives to engage in practices that make resources scarcer to capture additional rents.

The cost-alleviation effect operates through (i) complementarity with strong content quality and (ii) the fact that congestion  $\omega$  is a negative function of the quality of the contracting content  $q_2$ . For (ii), given the level of content quality  $q$  and network capacity  $\mu$ , the congestion levels are respectively given by

$$\omega_n(q, \mu) = \frac{\Delta\gamma_+}{2\mu} + \frac{\theta\Delta q}{6t\mu}\Delta\gamma,$$

and

$$\omega_z(q, \mu) = \omega_n(q, \mu) - \frac{\mu\theta\Delta q + \Delta\gamma}{12t\mu^2}\Delta\gamma.$$

Hence, readily observed is that CPs have opposite effects on the congestion under the two regulatory regimes. In particular, investments in quality by the strong CP clearly increase the average congestion in the network, that is,  $\frac{\partial\omega_k(q^k, \mu_k)}{\partial q_1^k} > 0$ , whereas investment by the weak CP always reduces it, that is,  $\frac{\partial\omega_k(q^k, \mu_k)}{\partial q_2^k} < 0$ .

Next, for the first component (i), by differentiating the expressions of the optimal content quality derived in lemma 4 with respect to  $\mu_z$ , we find that a marginal decrease in network capacity actually decreases the content quality offered by the non-zero-rated CP and increases the quality of the zero-rated CP<sup>14</sup>. Then, given this relationship with respect to congestion levels, a marginal decrease in network capacity also implies a reduction in network congestion. Finally, from the comparative statics on  $\mu_z^*$ , one can observe that the profit-maximizing ISP reduces its investment accordingly in response to a decrease in the non-zero-rated CP's quality, whereas the optimal network capacity is increasing in the zero-rated CP's quality provision<sup>15</sup>. Finally, recall that lemma 4 indicates that zero-rating implies a greater downward quality distortion from the non-zero-rated CP than from its competitor. As a result, the effect of this quality distortion on the level of  $\mu_z^*$  outweighs the impact of the weak content's quality degradation, and the ISP is not incentivized to invest in broadband capacity overall because quality distortion reduces its management costs.

## 5 Welfare analysis

In this section, we consider the welfare implications of allowing the ISP to depart from a strict net neutrality regulation. We first analyze the extent to which the ISP's optimal pricing policy differs from the social optimum and find that, unsurprisingly, a profit-maximizing ISP always

<sup>14</sup>Indeed, one obtains  $\partial q_1^{z^*} / \partial \mu_z = \frac{1}{\mu_z} \mathcal{A}_1(\theta, t, \mu_z) > 0$  and  $\partial q_2^{z^*} / \partial \mu_z = \frac{1}{\mu_z} \mathcal{A}_2(\theta, t, \mu_z) < 0$ .

<sup>15</sup>Assuming that the ISP's program is well defined and that the second-order condition is verified, one obtains that  $d\mu_z^* / dq_1^z > 0$  and  $d\mu_z^* / dq_2^z < 0$ .

fails to adopt the optimal policy. We then compare the welfare-maximizing content quality and broadband capacity provision with private incentives to invest for CPs and the ISP.

**Socially optimal pricing policy.** Let us first consider the pricing policy of the ISP with respect to the socially optimal policy and assume that the regulator was to set the per-unit fee imposed by the ISP on the use of the non-zero-rated content. Consider the optimal overage fee  $\tau^{\mathcal{W}}(\cdot)$  that, given content qualities  $\mathbf{q}_z = (q_1^z, q_2^z)$ , maximizes welfare as defined by

$$\mathcal{W}_k(\mathbf{q}_k, \mu_k) \equiv CS_k + \sum_i \pi_i^k + \Pi_k^{\text{ISP}},$$

for  $k = z, n$ . We obtain that the welfare-maximizing overage fee  $\tau^{\mathcal{W}}(\cdot)$  equals

$$\tau^{\mathcal{W}}(\mathbf{q}_z) = 2\theta(q_2^z - q_1^z),$$

leading to subgame symmetric equilibrium content qualities  $q_1^z = q_2^z = \theta$ . Hence, the optimal tariff that maximizes the social surplus has a per-unit surcharge equal to zero at the symmetric outcome, and we conclude the following:

**Proposition 5.** *A profit-maximizing ISP sets a per-unit fee  $\tau^*$  greater than the welfare-maximizing two-part tariff, whereas a profit-maximizing ISP sets the socially optimal policy, that is,  $\tau^{\mathcal{W}} = \tau^* = 0$ , if no congestion externalities exist.*

This result echoes the standard results on optimal two-part tariffs with homogenous consumers; that is, the optimal pricing policy for the ISP has a per-unit fee equal to the marginal cost, which is assumed to be zero. However, in the presence of congestion, because the surcharge acts as an excise tax, the quantity traded in equilibrium is reduced such that a profit-maximizing ISP fails to adopt the welfare-maximizing policy.

**Welfare implications.** We now state the main result of this section. Regarding the total surplus, consumer surplus, and industry profits, we find that consumers are always better off under a strict regulatory regime, whereas the total surplus is higher under net neutrality than zero-rating over some interval of network capacities. In either case, we find that the content sector's profits are always higher under zero-rating than net neutrality regulation.

**Proposition 6.** *(i) In a symmetric equilibrium, allowing a profit-maximizing ISP to financially discriminate between contents by setting a two-part tariff with a positive per-unit surcharge is welfare reducing for all  $\mu > \mu^w(\theta, t)$ . (ii) The content sector has higher profits under zero-rating, whereas consumers are better off under net neutrality irrespective of network capacity.*

*Proof.* If content quality is exogenous, we find that  $W_n(q_n, \mu_n) > W_z(q_z, \mu_z)$ , whereas this relation is unclear whenever we account for the equilibrium content quality provision. Computing the difference yields

$$\mathcal{W}_n(\mathbf{q}_n^*, \mu_n) - \mathcal{W}_z(\mathbf{q}_z^*, \mu_z) = \frac{1}{12} \left( \theta^2 + \left( \frac{1}{\mu_z} - \frac{1}{\mu_n} \right) 6\Delta_\gamma + \frac{3(9t - 5\theta^2)\Delta_\gamma^2}{(\theta^2 - 18t)^2 \mu_z^2} \right) + C(\mu_z) - C(\mu_n),$$

which for a given network capacity reduces to

$$\mathcal{W}_n(\mathbf{q}_n^*, \mu) - \mathcal{W}_z(\mathbf{q}_z^*, \mu) = \theta^2 + \frac{3(\gamma_1 - \gamma_2)^2(9t - 5\theta^2)}{(\theta^2 - 18t)^2 \mu^2}.$$

This quantity is strictly positive if

$$\mu > \sqrt{3} \left( \frac{(5\theta^2 - 9t)}{\theta^2(\theta^2 - 18t)^2} \right)^{1/2} (\gamma_1 - \gamma_2) \equiv \mu^w(\theta, t),$$

which is satisfied for  $\theta^2 < 9t < 5\theta^2$ .

Next, with respect to content industry profits, one obtains that

$$\Pi_n^{\text{CPs}}(\mathbf{q}_n^*) - \Pi_z^{\text{CPs}}(\mathbf{q}_z^*, \mu_z) = \frac{(\gamma_1 - \gamma_2)^2(\theta^2 - 36t)}{4(\theta^2 - 18t)^2 \mu_z^2} - \frac{\theta^2}{12},$$

which is negative under assumption **A2**.

Finally, computing the difference  $\text{CS}_n(\mathbf{q}_n^*) - \text{CS}_z(\mathbf{q}_z^*, \mu_z)$  given equilibrium content qualities  $\mathbf{q}_k^* = (q_i^{k*}, q_j^{k*})$  yields

$$\text{CS}_n(\mathbf{q}_n^*) - \text{CS}_z(\mathbf{q}_z^*, \mu_z) = \frac{3(\gamma_1 - \gamma_2)^2(33t - 2\theta^2)}{4(\theta^2 - 18t)^2 \mu_z^2} > 0,$$

and we obtain that  $\text{CS}_n(\mathbf{q}_n^*) > \text{CS}_z(\mathbf{q}_z^*, \mu_z)$ , and the assertions in the proposition follow.  $\square$

Although users enjoy a lower access price under zero-rating agreements, as  $H_n > H_z$ , they are subject to more advertising, that is,  $a_z^* > a_n^*$ , which supports an overall decline in content quality because we have shown that zero-rating softens CPs' quality competition in the content market, that is,  $q_i^n > q_i^z$ <sup>16</sup>. As a result, consumer surplus is greater under net neutrality regulations. Moreover, the content industry's profits are strictly higher, and we obtain that the result highlighted by proposition 1, which is that zero-rating actually increases industry profits with symmetric content qualities, carries over to optimal content quality. Next, in light of proposition 4, for  $\mu > \mu^w(\theta, t)$ , the ISP is shown to have incentives to implement zero-rating. Hence, allowing zero-rating agreements implies pervasive incentives from the ISP. As given in proposition 4, the ISP has incentives to engage in practices that make resources scarcer and finds it profitable to financially discriminate between CPs despite welfare reduction. The loss in consumer surplus implied by the profit-maximizing ISP's policy is not offset by the gain in profits realized in the content sector; thus, welfare is lower in a discriminatory network, and consumers are better off under a neutral regime.

<sup>16</sup>To see that the access price is lower under zero-rating, computing the difference  $H_n - H_z$  yields

$$H_n - H_z = \frac{(\gamma_1 - \gamma_2)(3t\mu - \gamma_1 + \gamma_2)}{12\mu^2 t} > 0$$

with exogenous qualities if  $\mu > \frac{\gamma_1 - \gamma_2}{3t}$  and

$$H_n - H_z = -\frac{3(\gamma_1 - \gamma_2)(3(\gamma_1 - \gamma_2) + \mu(\theta^2 - 9t)) + \theta^2 \mu^2(\theta^2 - 18t)}{6\mu^2(18t - \theta^2)} > 0$$

with optimal qualities if  $\mu > \bar{\mu}(\theta, t) = \frac{3\Delta\gamma}{9t - \theta^2}$ .

We conclude this section by considering as corollaries the efficient levels of content qualities and broadband provision by comparing private and social incentives to invest. We first derive the socially optimal level of content qualities and find that, in contrast to private incentives, the noncontracting CP should overinvest under zero-rating with respect to its contracting competitor, which should provide lower quality than it would in a neutral network. Next, when deriving the optimal level of broadband capacity, we find that a profit-maximizing pricing policy entails the underprovision of capacity whenever zero-rating is implemented.

**Efficient content quality provision.** Let us compare the quality provision under both regimes  $(q_n^*, q_z^*)$  with the socially optimal levels of quality provision  $(q_n^{\mathcal{W}}, q_z^{\mathcal{W}})$ , which is the solution of

$$(q_{1,k}^{\mathcal{W}}, q_{2,k}^{\mathcal{W}}) \in \underset{q_1^k, q_2^k}{\operatorname{argmax}} CS_k + \sum_i \pi_i^k + \Pi_k^{\text{ISP}}, \text{ for } k = n, z.$$

Computing and solving for the first-order condition for  $k = n, z$  yields, for  $i, j = 1, 2$  and  $i \neq j$ ,

$$\begin{aligned} q_{i,n}^{\mathcal{W}} &= \frac{\theta}{2}, \\ q_{i,z}^{\mathcal{W}} &= q_{i,n}^{\mathcal{W}} - \mathcal{D}_i(\theta, t, \mu_z), \end{aligned}$$

where  $\mathcal{D}_i(\theta, t, \mu_z) \equiv \frac{5\theta\Delta_{\gamma}}{2\mu_z(36t-11\theta^2)}$  and  $\Delta_{\gamma_i} = \gamma_i - \gamma_j$ .

It is immediate that  $q_i^{n*} < q_{i,n}^{\mathcal{W}}$ , such that CPs underprovide quality in the content market with respect to the socially optimal level under net neutrality. Given assumption A2, we have that  $\mathcal{D}_1(\theta, t, \mu_z) < 0$  for CP  $i = 1$  and  $\mathcal{D}_2(\theta, t, \mu_z) > 0$  for CP  $i = 2$ . Therefore,  $q_{1,z}^{\mathcal{W}} > q_n^{\mathcal{W}} > q_{2,z}^{\mathcal{W}}$ . In contrast to private incentives, the socially optimal level of quality for the noncontracting CP is greater than under net neutrality and greater than the zero-rated content quality. Because we have also shown that  $q_1^{z*} < q_{1,z}^{\mathcal{W}}$ , allowing for the implementation of a zero-rating contract implies a high degree of quality distortion of the noncontracting content and a lower distortion for the contracted content with respect to the optimum level of quality provided by the regulator.

**Efficient broadband capacity provision.** Let us consider the efficient level of broadband provision by a profit-maximizing ISP. Opponents of net neutrality regulation argue that imposing strict net neutrality results in the underprovision of resources by a profit-maximizing ISP with respect to the socially optimal level of network capacity. In the corollary that follows, we show that this is not true in our present model because the ISP provides capacity at the socially optimal level in the neutral network, whereas allowing for a discriminatory regime would result in underprovision with respect to the social optimum. Indeed, incentives to invest are given by a marginal change in profits following a marginal change in network capacity. Thus, we compare private and regulator incentives through the difference  $\frac{\partial \mathcal{W}_k}{\partial \mu_k} - \frac{\partial \Pi_k^{\text{ISP}}}{\partial \mu_k}$ . For the neutral situation, case (ii), the difference yields  $\frac{\partial \mathcal{W}_n}{\partial \mu_n} - \frac{\partial \Pi_n^{\text{ISP}}}{\partial \mu_n} = 0$ . For the discriminatory regime, case (iii), the difference leads to the following relation:

$$\frac{\partial \mathcal{W}_z}{\partial \mu_z} - \frac{\partial \Pi_z^{\text{ISP}}}{\partial \mu_z} = \frac{(\gamma_1 - \gamma_2)^2 (63t - 5\theta^2)}{2(\theta^2 - 18t)^2 \mu_z^3} > 0,$$

under assumption A2. We can then conclude this section with the following corollary:

**Corollary 2.** (i) *CPs' investments are suboptimal, and there is underprovisioning in content quality.* (ii) *A profit-maximizing ISP's investment in broadband capacity under net neutrality is socially optimal, whereas (iii) the regulator has greater incentives than does the ISP to invest in broadband capacity under zero-rating.*

In contrast to the work by [Gautier and Somogyi \(2020\)](#) for which the level of investment by an ISP is socially optimal, these results show that with explicit congestion and endogenous content qualities, a profit-maximizing ISP underprovides broadband capacity in equilibrium with respect to the socially optimal level. It has incentives to maintain scarce network capacity to extract more revenues from users of non-zero-rated content and makes CPs bear the cost of managing network congestion.

## 6 Discussion

In this section, we first discuss the two effects that we found regarding the incentives to invest in network capacity. In a second step, we relax the fixed demand hypothesis, considering that the zero-rating contract can increase the demand intensity of users consuming zero-rated content. We discuss the effects on the ISP's pricing policy and the total surplus generated in a discriminatory regime compared with a net neutrality regime.

**Incentives to invest in network capacity by a profit-maximizing ISP.** [Choi and Kim \(2010\)](#) isolated two effects that influence ISPs' incentives to invest in network capacity: a rent extraction effect and a network access charge effect. Investment in capacity affects how the ISP can extract revenue from end-users through access charges and from content providers through the priority line price. In our model, content providers do not pay to be zero-rated, but both effects are present on the user side. Nevertheless, given the interaction between congestion and content quality/functionality, we obtain an additional indirect effect through the strategic interaction between the ISP and content providers. We can decompose the marginal effect of an increase in network capacity for net neutrality as follows:

$$\frac{\partial \Pi_n^{\text{ISP}}}{\partial \mu} = \frac{\partial H_n}{\partial \mu} = \frac{\gamma_1 + \gamma_2}{2\mu^2},$$

and for a discriminatory regime as follows:

$$\frac{\partial \Pi_z^{\text{ISP}}}{\partial \mu} = \frac{\partial H_z}{\partial \mu} + \frac{\partial}{\partial \mu} (\tau x_z) = (\theta - a_2^z) \frac{\partial q_2^z}{\partial \mu} + \frac{1}{\mu} \left( \omega_z - (\gamma_1 - \gamma_2) \frac{\partial x_z}{\partial \mu} \right) + t \frac{\partial x_z}{\partial \mu} + \frac{\partial}{\partial \mu} (\tau x_z).$$

Computing the difference, we denote by  $\Delta_\Pi \equiv \frac{\partial \Pi_z^{\text{ISP}}}{\partial \mu} - \frac{\partial \Pi_n^{\text{ISP}}}{\partial \mu}$ , yields,

$$\Delta_\Pi = \underbrace{\left[ \underbrace{(\theta - a_2^z) \frac{\partial q_2^z}{\partial \mu}}_{<0} + \underbrace{\frac{\gamma_1 - \gamma_2}{\mu} \left( \frac{1}{\mu} \left( x_z - \frac{1}{2} \right) - \frac{\partial x_z}{\partial \mu} \right)}_{<0} + \underbrace{t \frac{\partial x_z}{\partial \mu}}_{>0} \right]}_{<0} + \underbrace{\tau \frac{\partial x_z}{\partial \mu} + \frac{\partial \tau}{\partial \mu} x_z}_{<0}.$$

Note that, as in [Choi and Kim \(2010\)](#), increasing network capacity under net neutrality allows the ISP to charge more users because the marginal user is not affected. However, the effect of capacity expansion is less clear-cut under a discriminatory regime because it affects the location of the marginal user and the marginal willingness to pay through the quality of the content being zero-rated. As in [Choi and Kim \(2010\)](#), we observe a similar term,  $\frac{1}{\mu} \left( \omega_z - (\gamma_1 - \gamma_2) \frac{\partial x_z}{\partial \mu} \right)$ , that reflects the effect of capacity expansion through market shares and congestion. In contrast to their observation with paid prioritization, under zero-rating capacity expansion induces a demand expansion for the non-zero-rated CP, that is,  $\frac{\partial x_z}{\partial \mu} > 0$ , which increases congestion because  $\omega_z$  is negatively biased toward this CP and reduces the willingness to pay for using the non-zero-rated content. The rent extraction effect in [Choi and Kim \(2010\)](#) concerns the ability of the ISP to extract rent from the CPs when selling the priority channel. Increasing network capacity reduces the value of the first priority. In our model, the ISP extracts rent from users who value non-zero-rated content through a higher unit surcharge. More precisely, capacity expansion reduces the rent of the ISP from the delivery of the non-zero-rated content, which further weakens its incentives to invest.

In our model, a close interaction exists between the ISP's pricing rule and investment strategy and the CPs' investment decision, both of which affect users' marginal willingness to pay. This additional effect, which is not present in [Choi and Kim \(2010\)](#), goes through a quality adjustment of the content offered by the zero-rated CP and is captured by the term  $(\theta - a_2^z) \frac{\partial q_2^z}{\partial \mu} < 0$ . We observe that by investing in capacity, the ISP negatively affects access charges by reducing the quality of the zero-rated content, which further reduces the marginal willingness to pay to access the non-zero content. Finally, we derive another indirect channel related to the previous effect highlighted, which does not readily appear in the marginal profits comparison. To the extent that investments in content quality and investments in network capacity interact, investments in network capacity by the ISP and in quality by the CPs are strategic complements between the non-zero-rated CP and the ISP, that is,  $\frac{\partial^2 \pi_1}{\partial q_1 \partial \mu} > 0$ ,  $\frac{\partial^2 \Pi^{ISP}}{\partial \mu \partial q_1} > 0$ , and strategic substitutes between the zero-rated CP and the ISP, that is,  $\frac{\partial^2 \pi_2}{\partial q_2 \partial \mu} < 0$ ,  $\frac{\partial^2 \Pi^{ISP}}{\partial \mu \partial q_2} < 0$ . As already mentioned in the text, because of this interaction, downward quality distortion implies a reduction in network congestion, which has two effects. The first is that a reduction in congestion actually increases users' marginal willingness to pay for non-zero-rated content, as shown by the marginal profits previously presented. Therefore, because the quality distortion is greater for non-zero-rated content, strategic complementarity with network capacity further weakens the incentives to invest. The second effect is that because reducing capacity leads to a downward distortion of quality, which reduces congestion, as [lemma 4](#) goes, the mere existence of the interaction between quality and capacity leads to savings in network management costs for the ISP, which reduces the need for capacity investment.

**Assuming variable demand for zero-rated content.** Given the fixed demand assumption, the negative welfare result is mainly from a decrease in content quality and an increase in advertising per unit (increase in advertising disutility). We approach the discussion with exogenous content qualities because, unfortunately, the model is not tractable when the quality is endogenous. To do this, let us assume that users consume an additional unit of zero-rated content. This consumption gives a benefit of  $\eta$  to the user, and we can assume that a propor-

tion  $r$  of users consume an extra unit, whereas a proportion  $1 - r$  does not. Then, the expected utility of users who can choose one or two units of the zero-rated content is given by:

$$U_i = \theta q_i^k - t|x - l_i| - \omega^k(\mathbf{x}, \mu_k) - q_i^k a_i^k - H_z - \tau,$$

whereas the expected utility of users who consume the zero-rated content is given by:

$$U_i = r \left( \theta q_i^k - t|x - l_i| - \omega^k(\mathbf{x}, \mu_k) - q_i^k a_i^k - H_z + \eta \right) \\ + (1 - r) \left( \theta q_i^k - t|x - l_i| - \omega^k(\mathbf{x}, \mu_k) - q_i^k a_i^k - H_z \right),$$

which reduces to

$$U_i = \theta q_i^k - t|x - l_i| - \omega^k(\mathbf{x}, \mu_k) - q_i^k a_i^k - H_z + \eta r.$$

Let us define the following relation when content provider  $i = 1, 2$  is zero-rated:

$$\mathcal{B}(\theta, \mathbf{q}^{ZRi}, \mathbf{a}^{ZRi}) \equiv \left( \theta - a_1^{ZRi} \right) q_1^{ZRi} - \left( \theta - a_2^{ZRi} \right) q_2^{ZRi}.$$

If the strong content provider (CP 1) is zero-rated, the marginal users are given by

$$\hat{x}^{ZR1} = \frac{1}{2} + \frac{1}{2t} (\tau + \eta r) + \frac{1}{2t} \mathcal{B}(\theta, \mathbf{q}^{ZR1}, \mathbf{a}^{ZR1}) = \hat{x}_z^{Fixed} + \frac{\eta r}{2t},$$

which implies total traffic  $D^{ZR1} = 1 + r\hat{x}^{ZR1}$  and yields congestion level,

$$\omega^{ZR1} = \frac{((1+r)\gamma_1 - \gamma_2)\hat{x}^{ZR1} + \frac{\gamma_2}{\mu_z}}{\mu_z}.$$

If the weak content provider (CP2) is zero-rated, the marginal users are given by

$$\hat{x}^{ZR2} = \frac{1}{2} - \frac{1}{2t} (\tau + \eta r) + \frac{1}{2t} \mathcal{B}(\theta, \mathbf{q}^{ZR2}, \mathbf{a}^{ZR2}) = \hat{x}_z^{Fixed} - \frac{\eta r}{2t},$$

which implies total traffic  $D^{ZR2} = 1 + r(1 - \hat{x}^{ZR2})$  and yields congestion level,

$$\omega^{ZR2} = \frac{(\gamma_1 - (1+r)\gamma_2)\hat{x}^{ZR2} + \frac{(1+r)\gamma_2}{\mu_z}}{\mu_z}.$$

It is then immediate to observe that zero-rating the strong content provider implies more traffic and congestion on the network. As subsequently shown, the ISP finds it optimal to zero-rate the weak content provider because the per-unit surcharge is otherwise always negative. The per-unit fee associated with zero-rating the strong CP is given by

$$\tau^{ZR1} = -\frac{(1+r)\gamma_1 - \gamma_2}{2\mu_z} - \eta r < 0.$$

The per-unit fee associated with the weak CP being zero-rated is given by

$$\tau^{ZR2} = \frac{\gamma_1 - (1+r)\gamma_2}{2\mu_z} - \eta r,$$

from which we immediately observe that higher intensity of demand in the weak content results in lower per-unit surcharge collected in relation to the strong content, that is,  $\partial \tau^{ZR2} / \partial r <$

0. Using the previous expressions, we found that the difference in the degree of advertisement exposition between content providers is lower with flexible demand. More precisely, the non-zero-rated CP increases advertising with the intensity of demand, whereas the zero-rated CP decreases the level of advertising to which it exposes its users. For exogenous content qualities, the optimal advertising exposure rates are given by

$$a_1^{ZR2} = a_1^{Fixed} + \frac{\gamma_2 - \eta\mu_z}{6q_z\mu_z}r \quad \text{and} \quad a_2^{ZR2} = a_2^{Fixed} - \frac{\gamma_2 - \eta\mu_z}{6q_z\mu_z}r,$$

such that the result obtained in equation (12) for both content providers remains qualitatively the same and is robust to the introduction of flexible demand.

Next, we argue that increasing demand intensity through zero-rating is optimal for the ISP only if it is coupled with a reduction in network capacity such that the qualitative result obtained in proposition 4, which states that a switch to zero-rating implies lower investment in network capacity, is robust to flexible demand. Given flexible demand, a constraint exists on the proportion of users who intensify their demand for zero-rated content. The optimal per-unit fee is only positive when the proportion of users consuming more than one unit of the zero-rated content is such that

$$r < \frac{\gamma_1 - \gamma_2}{\gamma_2 + \eta\mu_z} \equiv \tilde{r}.$$

Therefore, the ISP will not apply zero-rating when  $r > \tilde{r}$  and, for a given network capacity, the greater the intensity of demand, the less willing the ISP is to switch to a discriminatory regime. The benefits that the ISP expects to derive from users who consume zero-rated content are not offset by the actual rent from users who consume non-zero-rated content. We note that the threshold  $\tilde{r}$  decreases as the network capacity increases. Therefore, traffic inflation should increase the ISP's incentive to create a resource shortage. The reduction in network capacity allows the ISP to capture an additional share of the residual rent left by the CPs through an increase in the per-unit surcharge to users. Therefore, as shown in the base model, switching to a discriminatory regime under flexible demand also reduces incentives to invest in network capacity.

Finally, we cannot derive a closed-form solution when comparing total welfare under net neutrality and total welfare under the discriminatory network with flexible demand. Nevertheless, we can show that, with exogenous content qualities, the result obtained in the basic model as derived in Proposition 6 is less sharp. For given content qualities, we find in the basic model that switching to a discriminatory regime always reduces welfare. With flexible demand, we observe that, depending on the degree of demand intensity, a discriminatory regime can either reduce or improve welfare. To this end, we set the parameters to  $\gamma_1 = 2/3$ ,  $\gamma_2 = 1/3$ ,  $t = 1$ , and  $\eta = 1$  and calculate the difference in total surplus between the two regulatory regimes, that is,  $\Delta_W \equiv \mathcal{W}_z - \mathcal{W}_n$ . This difference is represented in figure 1, in which the demand intensity  $r$  is on the  $x$ -axis and  $\Delta_W$  on the  $y$ -axis. While we cannot derive its exact value, we observe that there exists a threshold  $\hat{r} > 0$  in the demand intensity, which with  $\tilde{r} > 0$  guarantees the existence of parameter region so that it is profitable for the ISP to zero-rate the weak CP, and such that for any  $r < \min\{\hat{r}, \tilde{r}\}$ , a deviation from network neutrality is welfare-reducing, that is,  $\Delta_W < 0$ , and for any  $r \in [\hat{r}, \tilde{r}]$ , a discriminatory regime is welfare-enhancing, that is,  $\Delta_W > 0$ .



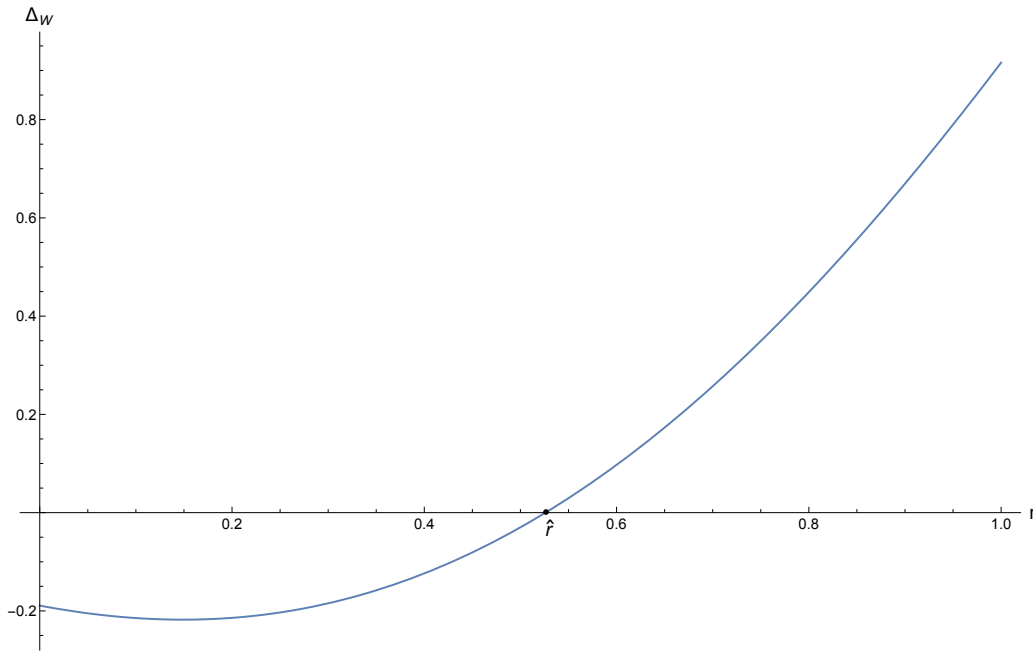


Figure 1: Difference in total welfare between zero-rating and net-neutrality.

## 7 Conclusion

This work contributes to the debate on net neutrality and possible departures from it by considering a profit-maximizing ISP that financially discriminates among CPs through exclusive zero-rating contracts. We study the implications for the quality of content provision and investments in broadband infrastructure, focusing on the advertising-supported business model employed in the content market.

In our model, a profit-maximizing ISP finds it profitable to deviate from net neutrality, although zero-rating is welfare-reducing when the degree of congestion is low and always harms consumers. We find that zero-rating hinders innovation at the sector level, which contradicts opponents of the main arguments for net neutrality regulation. However, individual incentives to provide quality investments are misaligned between the two CPs. Zero-rating makes content more expensive for consumers to access and imply a downward distortion of content quality by increasing downward vertical differentiation. Through zero-rating, the ISP is able to increase the CPs' asymmetry and affect consumers' willingness to pay for content. The CP that contracts under a zero-rating plan have much stronger incentives to innovate than that of its rival, which might purposely degrade the quality of its services to minimize its costs. Hence, CPs move from a minimal differentiation equilibrium to a downward vertical differentiation outcome. Next, we show that zero-rating possibly reduces congestion, whereas a profit-maximizing ISP always underinvests in broadband infrastructure in the discriminatory network. We highlight that this underprovision comes from a standard rent-extraction argument and that a new cost-alleviation channel is also at play, which relates to the complementarity between network capacity and content quality. As our results show, and as a complement to the current literature, the debate on net neutrality is far from straightforward, and the relation among zero-rating, content innovation, and broadband provision is not clear-cut. Although we attempt to capture salient features of the market for content provision, the

model is not without limitations.

First, we do not address the implications of paid zero-rating contracts by CPs, a prominent form of contracting in the mobile industry. In a future extension, we plan to introduce another stage in which the ISP sells zero-rating contracts through a sealed-bid, first-price auction, as is the case for paid prioritization offers (see [Choi and Kim, 2010](#), and [Gautier and Somogyi, 2020](#)). Introducing side payments from CPs to the ISP provides another channel for the ISP to extract surplus from CPs and might strengthen incentives to distort competition. Second, one of the concerns of policymakers regarding departures from net neutrality is that doing so might impose barriers to entry for new firms in the content provision sector. This issue represents an interesting extension and a possible avenue for future research. Third, at the consumer level, the data cap does not exist. New research could consider not only a non-zero data cap but also results for distinct levels of data caps to more closely capture the design of certain ISP offers. Fourth, we do not consider competition between ISPs, which could yield distinct results. Finally, we consider the total level of congestion on the entire network without interaction with the attractiveness of content and possible network effects. A possible channel for improvement is to consider an individual congestion-quality interaction to introduce indirect network effects into the model. Doing so would allow us to capture the idea that more attractive content results in more users consuming and taking the time to use it and increases relative network congestion.

## References

- Anderson, S. P., Foros, Ø. and Kind, H. J.: 2017, Product functionality, competition, and multipurchasing, *International Economic Review* **58**(1), 183–210.
- Bourreau, M., Kourandi, F. and Valletti, T.: 2015, Net neutrality with competing internet platforms, *The Journal of Industrial Economics* **63**(1), 30–73.
- Calzada, J. and Tselekounis, M.: 2018, Net neutrality in a hyperlinked internet economy, *International Journal of Industrial Organization* **59**, 190–221.
- Choi, J. P., Jeon, D.-S. and Kim, B.-C.: 2018, Net neutrality, network capacity, and innovation at the edges, *The Journal of Industrial Economics* **66**(1), 172–204.
- Choi, J. P. and Kim, B.-C.: 2010, Net neutrality and investment incentives, *The RAND Journal of Economics* **41**(3), 446–471.
- Easley, R. F., Guo, H. and Krämer, J.: 2018, Research commentary: From net neutrality to data neutrality: A techno-economic framework and research agenda, *Information Systems Research* **29**(2), 253–272.
- Economides, N. and Hermalin, B. E.: 2012, The economics of network neutrality, *The RAND Journal of Economics* **43**(4), 602–629.
- Economides, N. and Tåg, J.: 2012, Network neutrality on the internet: A two-sided market analysis, *Information Economics and Policy* **24**(2), 91–104.
- Gautier, A. and Somogyi, R.: 2020, Prioritization vs zero-rating: Discrimination on the internet, *International journal of industrial organization* **73**, 102–662.
- Goldfarb, A. and Tucker, C.: 2019, Digital economics, *Journal of Economic Literature* **57**(1), 3–43.

- Greenstein, S., Peitz, M. and Valletti, T.: 2016, Net neutrality: A fast lane to understanding the trade-offs, *Journal of Economic Perspectives* **30**(2), 127–50.
- Hermalin, B. E. and Katz, M. L.: 2007, The economics of product-line restrictions with an application to the network neutrality debate, *Information Economics and Policy* **19**(2), 215–248.
- Hoernig, S. and Monteiro, F.: 2020, Zero-rating and network effects, *Economics Letters* **186**, 108–813.
- Inceoglu, F. and Liu, X.: 2019, Multiproduct price discrimination with quantity limits: An application to zero-rating, *Economics Letters* **180**, 41–45.
- Jeitschko, T. D., Kim, S. J. and Yankelevich, A.: 2020, Zero-rating and vertical content foreclosure, *Information Economics and Policy* p. 100899.  
**URL:** <http://www.sciencedirect.com/science/article/pii/S0167624520301438>
- Jullien, B. and Sand-Zantman, W.: 2018, Internet regulation, two-sided pricing, and sponsored data, *International Journal of Industrial Organization* **58**, 31–62.
- Krämer, J. and Peitz, M.: 2018, A fresh look at zero-rating, *Telecommunications Policy* **42**(7), 501 – 513.
- Krämer, J. and Wiewiorra, L.: 2012, Network neutrality and congestion sensitive content providers: Implications for content variety, broadband investment, and regulation, *Information Systems Research* **23**(4), 1303–1321.
- Lee, R. S. and Wu, T.: 2009, Subsidizing creativity through network design: Zero-pricing and net neutrality, *Journal of Economic Perspectives* **23**(3), 61–76.
- Malone, J., Nevo, A. and Williams, J. W.: 2017, The tragedy of the last mile: Economic solutions to congestion in broadband networks, *Available at SSRN 2849869* .
- Nevo, A., Turner, J. L. and Williams, J. W.: 2016, Usage-based pricing and demand for residential broadband, *Econometrica* **84**(2), 411–443.
- Nurski, L.: 2012, Net neutrality, foreclosure and the fast lane: An empirical study of the uk.
- Peitz, M. and Schuett, F.: 2016, Net neutrality and inflation of traffic, *International Journal of Industrial Organization* **46**, 16–62.
- Reggiani, C. and Valletti, T.: 2016, Net neutrality and innovation at the core and at the edge, *International Journal of Industrial Organization* **45**, 16–27.
- Schnurr, D. and Wiewiorra, L.: 2018, Bit-by-bit towards unlimited: An analysis of zero rating and sponsored data practices of internet service providers, *29th European Regional ITS Conference, Trento 2018*, number 184965, International Telecommunications Society (ITS).
- Somogyi, R.: 2017, The economics of zero-rating and net neutrality, *Unpublished manuscript. Université catholique de Louvain, Center for Operations Research and Econometrics (CORE)* .