GovReg
WORKING PAPER SERIES

Zero Rating, Content Quality and Network Capacity

Emmanuel Lorenzon

December 9, 2020

RESEARCH
Dauphine | PSL
CHAIRE GOUVERNANCE ET RÉGULATION

Zero Rating, Content Quality and Network Capacity*

Emmanuel Lorenzon[1,2]

[1]Univiversity Grenoble Alpes, CNRS, INRAE, Grenoble INP, GAEL, 38000 Grenoble, France
[2]Governance and Regulation Chair, University Paris-Dauphine, PSL Research University, 75016 Paris, France

December 9, 2020

Abstract

We consider a departure from net neutrality by an Internet service provider (ISP) that financially discriminates among content providers through bilateral zero rating contracts. Zero rating is an instrument to distort competition between content providers and the way 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 access and implies a downward distortion of quality by increasing downward vertical differentiation. Content providers move from a minimal differentiation equilibrium to a downward vertical differentiation outcome. Next, we find that while zero rating happens to reduce congestion, a profit-maximizing ISP always underinvests in the broadband infrastructure in the discriminatory network. We highlight that this underprovision comes from a standard rent-extraction argument and a new cost-alleviation channel, which relates to the complementarity between network capacity and content quality. Finally, the ISP always implements zero rating, which is welfare reducing and detrimental to consumers.

*__Acknowledgments:__ I am grateful to Petros Sekeris, Carlo Cambini, Vicente Lagos, Axel Gautier, Wilfried Sand-Zantman, Steffen Hoernig, Marc Bourreau, Alexandre de CorniÃÍre, Yassine Lefouili and Byung-Cheol Kim for their very helpful comments and discussions. I am also grateful for the comments received from the participants at: the MaCCI 2019, the 11th Conference on Digital Economics, the EARIE 2019 and the JEI 2019.

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 over 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[1]. 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 Telenos, which has been found to use zero rating as a traffic management tool to discriminate against applications that were not part of its subscription plans MyChat and MyMusic[2].

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, and for all data packets consumed in excess of that allowance, the consumers are charged marginal fees and/or the usage of data is either blocked or restricted. Zero rating is thus a tool an ISP can implement to price discriminate among content providers (CPs). The data from zero-rated content do not count against the cap, and once a consumer reaches his or her data cap, such content is exempted from per-unit surcharges and usage restrictions[3].

Network operators advertise zero rating as beneficial for users, as the latter are purportedly able to consume more content paying the same price for their mobile plan and as it would allow operators to efficiently manage traffic and foster their incentives to invest in the quality of the network[4]. As CPs typically rely on traffic to generate revenue, either from advertisements or user payments, opponents of such data management regimes contend that departures from net neutrality might steer consumers' choices of online content towards providers included in the operator's contracting offers, which may raise barriers to entry and impede incentives to provide high content quality in the market for content[5]. While 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 been overlooked by the network neutrality literature with regard to zero rating practice[6]. However, there exists a tight relationship 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 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 rating is thought to be a tool to enhance content differentiation and to alter competition in the market for content.

[1] 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, as 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`.

[2] The judgement of the court is available at this link.

[3] 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.

[4] See Schnurr and Wiewiorra (2018) and Krämer and Peitz (2018).

[5] See Nurski (2012).

[6] See Krämer and Peitz (2018).

However, the focus has been on the ability of an ISP to mitigate content asymmetry in environments where CPs are passive and network investments are left separate[7]. CPs do not invest in their content's attractiveness, nor do they determine the amount of advertisement to which to expose their users, and the ISP faces some given capacity constraint. 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 focus on and compare two regulatory regimes: net neutrality and bilateral zero rating contracts[8]. To study the impact that a move from the net neutrality regime has 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 may price discriminate between CPs by charging consumers different per-unit fees for the two offerings of content. 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 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, which, as in Calzada and Tselekounis (2018), interacts with the quality of content. The idea is that the higher the quality of content is, the more time consumers spend using it, the more they are exposed to advertisements and the more CPs benefit from quality improvement. Hence, the per-unit advertising exposure rate is understood as some 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 competition and advertisement competition between CPs but also captures the business models of most services included in zero rating[9].

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

[7] See, for instance, Gautier and Somogyi (2020), Jeitschko et al. (2019) or Jullien and Sand-Zantman (2018).

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

[9] 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 where revenue is generated from advertising displayed to consumers using the service for free while paying consumers have access to advertising-free content.

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 is contracting with the ISP, benefits from a greater market share and provides a higher level of 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 non-contracting 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, so the content sector is 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. As 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 standard price-discrimination argument, which implies that the ISP contracts with the least attractive CP, zero rating reduces the ISP's incentives to invest in network capacity through a cost-alleviation channel. Indeed, in contradiction with the main claim from the industry, we find that a profit-maximizing ISP underinvests under a discriminatory regime. However, due to complementarity with content quality, we demonstrate that zero rating happens to reduce congestion on the network, which might constitute an instrument for traffic management. This underinvestment is motivated by two explanatory channels. The first operates via standard price discrimination: the ISP increases resource scarcity, which allows for a higher per-unit surcharge on consumers of non-zero-rated content. This implies that in equilibrium, as in Gautier and Somogyi (2020), the ISP contracts with the least attractive CP. Reducing network capacity allows for the extraction of a larger rent from users of the most attractive content. The second channel operates via by cost-reducing incentives: indirectly, the quality of the non-contracted content and network capacity are complements. By imposing a per-unit fee, the strong CP reduces the quality it provides to its users, which also reduces congestion on the network and alleviates the need for the ISP to invest in the network, and 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 with 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, from the social welfare analysis, we find 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 resources scarcer and always finds it profitable to financially discriminate between CPs, while 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 large literature on net neutrality and data management practices by ISPs that violate the principle[10]. In particular, it 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[11].

As underlined by Goldfarb and Tucker (2019), the literature on net neutrality developed with the technologies of data transmission and the growing importance of the role of ISP strategies for the emergence of other businesses. Lee and Wu (2009) consider 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 would 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, which are possible because CPs do not need to negotiate with distinct ISPs for access to their consumers, 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, as the latter would be able to consume more content while paying the same price for their mobile plan.

In a deregulated market, ISPs would have incentives to depart from net neutrality, as they can generate additional revenues from CPs by offering benefits in return (e.g., prioritization of data or exemptions from the users' data allowance), attract new customers due to network effects and better discriminate among consumers on price and quality[12]. Schnurr and Wiewiorra (2018) find that laissez-faire practices might distort competition towards the CPs included in the sponsorship plan and may 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 UK. The author finds that departures from net neutrality might steer consumers' choice towards the CP included in the zero rating plan. Once CPs typically rely on traffic-based revenues (advertising or consumer payment), a 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, as we find that a discriminatory regime reduces content quality and implies downward vertical differentiation. Due to the interplay among content quality, congestion and network capacity, zero rating reduces managerial costs, increases CPs' asymmetry and allows the ISP to collect greater payments from users of the non-contracting 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) note that different services require distinct network capacities and that departing from net neutrality

[10] See Easley et al. (2018) and Greenstein et al. (2016) for a survey.

[11] 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.

[12] See Krämer and Peitz (2018) and Goldfarb and Tucker (2019).

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 decision 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. While 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[13]. While 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 non-contracting CP due to the positive relationship between content quality and congestion and to 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) considers 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, and when content is attractive, the ISP always offers an open zero rating. Jullien and Sand-Zantman (2018) consider 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, as 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 CPs' value. While 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 and the composition of consumption. Zero rating is found to be welfare enhancing and to cause network capacity expansion. Jeitschko et al. (2019) consider 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 analyze two groups of consumers who are 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 may harm consumers. Gautier and Somogyi (2020) compare the market outcomes under both zero rating and paid prioritization with two CPs that are horizontally

[13] In Choi and Kim (2010), the rationale follows from the fact that the 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).

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 towards zero rating. Finally, Hoernig and Monteiro (2020) study the role of network effects in an ISP's rationale for implementing zero rating. They note 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 note that the result is similar under monopoly and duopoly, but 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 to 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, while 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 regulation (denoted n hereafter), and a discriminatory network, which stands represents zero rating bilateral contracting (denoted z hereafter).

2.1 Monopolistic ISP

Under both regimes, the ISP offers consumers a contract $\mathbf{\Phi} = (H, \kappa, \tau)$, which consists of a subscription fee (connection) H, a data allowance (data cap) κ and a uniform overage fee, τ, charged in excess of κ. In a discriminatory network, the ISP can propose a zero rating contract to one or several CPs. We then have that in a neutral regime, $\tau_i \equiv 0$, while under exclusive zero rating, we have that $\tau_i = 0$ and $\tau_j > 0$, when content i is zero rated, where j denotes non-zero-rated content.

The ISP provides a unique network capacity μ_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 μ_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

$$\Pi_k^{\text{ISP}} \equiv H_k + \tau_i D_i^k \mathbb{1}_{ZRj} - I(\mu_k)$$

with $\mathbb{1}_{d_j=z} = 1$ if CP j is zero-rated and 0 otherwise. In the 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.

2.2 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. The CPs are thus 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 during the time they use content, and the higher the quality of content is, the higher the ad exposure. This exposure to advertisements implies 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 $a_1 < a_2$, so 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

$$\pi_i^k(.) = q_i^k a_i^k D_i^k(.) - C(q_i), \tag{1}$$

where $D_i^k(.)$ denotes CP i's market share under regime $k = n, z$. Finally, we make the following standard assumptions with respect to the investment in quality cost function $C(.)$: $C'(.) > 0, C''(.) > 0$ 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) so that the quality exhibits the characteristics of a public good for consumers.

2.3 Consumers

The ISP charges consumers a connection fee H_n in the neutral network and a three-part tariff $\mathbf{\Phi} = (H_z, \kappa, \tau)$ in the discriminatory network. Each user consumes the content of only one CP, and a consumer of type $x \in [0, 1]$ that patronizes CP i has a net utility given by

$$U_i(.) \equiv \theta q_i^k - t|x - l_i| - \omega^k(\mathbf{x}, \mu_k) - q_i^k a_i^k - H_k - \max\{0, 1-\kappa\}\tau \mathbb{1}_{d_j = z}, \tag{2}$$

where θ denotes a gross surplus, which is assumed to be large enough, t is the transportation cost and κ is the data cap offered by the ISP that is assumed be set to $\kappa = 0$. By doing so, we only consider the effect of the overage fee τ. Firm i is allocated position l_i on the unit line such that CP 1 is at $l_1 = 0$ and CP 2 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.

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. We then define the average congestion as follows:

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

where $\left(n_1^e, n_2^e\right)$ are the expected market shares of the CPs. The level of congestion is a decreasing function of the network capacity μ and an increasing function of the content request rate γ_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, so that in equilibrium, the 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))

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

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

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 quality investment $\sum_{i=1,2} C(q_i)$ and the cost of network capacity investment $\phi(\mu)$, with

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

2.4 Timing

The timing of the game is as follows: at $t = 1$, the profit-maximizing ISP determines its investment level in congestion-reducing investment μ_k at cost $\phi(\mu_k)$. At $t = 2$, CPs simultaneously determine their investments level in quality q_i^k at cost $C(.)$. At $t = 3$, the ISP offers both contracts to CPs and chooses its optimal policy $(H_k, \tau) \in \mathbb{R}_+^2$. At $t = 4$, CPs either accept or reject the contract and simultaneously choose the advertising exposure rate a_i^k. Finally, at $t = 5$, consumers either subscribe or not and decide which 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 a more 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. In the following, we will make the following assumption:

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

3 Equilibrium analysis

In this section, we derive the equilibrium of each market configuration and then compare the implications that a departure from net neutrality has for market equilibrium values. We assume that content and network qualities are given and that the game is solved by backward induction. 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 surplus from the indifferent consumer. In the discriminatory network, which corresponds to a bilateral 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

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 γ_2, which we call the weak content provider.

Lemma 1. *Under the assumption that content request rates are such that $\gamma_1 > \gamma_2$, it is not profitable for the ISP to zero rate the CP with the highest content request rate*[14].

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, then the optimal amount of advertisement that CPs expose their users to 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 of

$$\theta q_1^k - t\hat{x}_k - \tau \mathbb{1}_{\mathrm{ZR}_2} = \theta q_2^k - t\left(1 - \hat{x}_k\right)$$

and is given by

$$\hat{x}_k = \frac{1}{2} + \frac{\theta\left(q_1^k - q_2^k\right) - a_1^k q_1^k + a_2^k q_2^k - \tau \mathbb{1}_{\mathrm{ZR}_j}}{2t}$$

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

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 each of their users. CP i's problem is then to find $a_i^{k^*}$ such that

$$a_i^{k^*} \in \underset{a_i^k > 0}{\operatorname{argmax}}\, q_i^k a_i^k D_i^k\left(\boldsymbol{a}^k, \boldsymbol{q}^k\right) - C\left(q_i^k\right),$$

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

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

so that it is clear that advertising exposure rates are strategic complements. Solving for both reaction functions yields the optimal amount,

$$a_1^{k^*} = \frac{\theta\left(q_1^k - q_2^k\right) + 3t - \tau \mathbb{1}_{\mathrm{ZR}_j}}{3q_1^k}\ ;\ a_2^{k^*} = \frac{\theta\left(q_2^k - q_1^k\right) + 3t + \tau \mathbb{1}_{\mathrm{ZR}_j}}{3q_2^k}. \tag{5}$$

Neutral network Given content qualities $\boldsymbol{q}_z^* = \left(q_1^{z*} q_2^{z*}\right)$, the ISP chooses its tariff $\left(H_n^*\right)$ to extract the marginal consumer's surplus, i.e., $\hat{U}\left(\hat{x}_n, H_n^*\right) = 0$, and sets the optimal connection fee to

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

[14] This follows from the restriction that we do not allow the ISP to charge a negative per-unit fee, $\tau < 0$

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:

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

and profits are:

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

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, i.e., $C(q_n) = 0$, 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 amount of advertising exposure as consumer heterogeneity increases, i.e., when t increases, and reduce this amount with content differentiation. Higher transportation costs enhance each firm's market power, which allows for higher exposure.

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

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

and

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

Assuming exogenous 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 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 with which its users are exposed to advertisement, while the zero-rated CP increases this amount accordingly because it benefits from the distortion in market shares. We then derive the following lemma:

Lemma 2. *Assuming symmetric content qualities, for a given per-unit fee, the more (less) content is vertically differentiated, the lower (higher) the exposure rate and the ad nuisance are for users.*

To observe the interaction between advertising exposure time and quality within the ISP's optimal pricing, let us assume that $\boldsymbol{a}^{z^*} = \left(a_1^{z^*}, a_2^{z^*}\right)$ and $\boldsymbol{q}^z = \left(q_1^z, q_2^z\right)$ are given. Again, the ISP extracts all the surplus from the marginal end user and sets $H_z(.)$ such that $\hat{U}\left(\hat{x}_z, H_z(.)\right) = 0$, which leads to the following:

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

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

$$\tau^* \in \underset{\tau}{\operatorname{argmax}} \; H_z\left(\boldsymbol{q}^z, \boldsymbol{a}^{z^*}, \tau\right) + \tau D_i^z\left(\boldsymbol{q}^z, \boldsymbol{a}^{z^*}, \tau\right),$$

which, given $\boldsymbol{a}^{z^*}$ and $\boldsymbol{q}^z$, leads to the following optimum overage fee[15],

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

where $\Delta_\gamma = \gamma_1 - \gamma_2$.

The optimum per-unit fee is increasing in the relative quality Δ_{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 a greater market power of each CP. This entails the possibility of charging consumers that purchase non-zero-rated content more.

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

$$\begin{aligned}(H_z^*, \tau^*) \in \underset{H_z, \tau>0}{\operatorname{argmax}}\ & H_z\left(\boldsymbol{q}^z, \boldsymbol{a}^{z^*}, \tau\right) + \tau D_i^z\left(\boldsymbol{q}^z, \boldsymbol{a}^{z^*}, \tau\right) \\ s.t \quad & \hat{U}\left(\hat{x}, H_z^*, \tau\right) = 0,\end{aligned}$$

which yields

$$H_z^* = \frac{\theta\Delta_{q_+^z}}{2} - \frac{\Delta_{\gamma_+}}{2\mu_z} - \frac{\theta\Delta_\gamma\Delta_{q^z}}{6t\mu_z} - \frac{3}{2}t + \frac{\Delta_\gamma}{6t\mu_z}\tau + \frac{\tau}{2} \tag{8}$$

$$\tau^*\left(\boldsymbol{q}^z, \mu_z\right) = \frac{\theta\Delta_{q^z}}{2} + \frac{\delta\Delta_\gamma}{2\mu_z}. \tag{9}$$

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

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

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

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

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

While under net neutrality, CPs share the market, it is clear that due to financially differentiated content, the profit-maximizing ISP is able to distort market competition through the difference in content requests and that this distortion is more likely to be high the lower the network capacity is.

[15] A sufficient condition for concavity of the profit function is $\theta\left(q_1^z - q_2^z\right) > 2\tau - \frac{\delta\Delta_\gamma}{\mu_z}$.

3.2 Effect on advertising exposure rates, access price 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, i.e., $q_i^k = \bar{q}$ with $k = n, z$. We begin by stating the main proposition we derive from this subsection, and we then consider each case in what follows.

Proposition 1. *Suppose that content quality is exogenous, i.e., $q_i^k = q_k$ 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 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, while it strictly reduces the non-contracting CP's profits.

Effect on advertising exposure. Part (i) of the proposition states that the optimum amount of 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

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

and

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

The optimum amount of 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, while the zero-rated CP increases the time exposure as this difference increases. In a neutral network, CPs share the market, i.e., $x_i^n = 1/2$, and expose consumers to the same amount of advertising. However, as the surcharge charged by the ISP acts as a standard excise tax, it is straightforward to observe that the non-zero-rated CP supports a reduction in its demand, while the zero-rated content supports an increase in its own, i.e., $x_1^z < x_1^n = x_2^n < x_2^z$. This 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, i.e., $a_2^z > a_1^z$. This makes sense, as 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 contrast, 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, as 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:

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

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

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

and

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

with $\Delta_\gamma = \gamma_1 - \gamma_2$ 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, and the greater the vertical differentiation and the 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 would be exposed.

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

$$H_n\left(q_n, \mu_n\right) = \theta q_n - \frac{3}{2}t - \frac{\Delta_{\gamma_+}}{2\mu_n},$$
$$H_z\left(q_z, \mu_z\right) = \theta q_z - \frac{3}{2}t - \frac{\Delta_{\gamma_+}}{4\mu_z} + \frac{\Delta_\gamma^2}{12t\mu_z} - \frac{1}{2\mu_z}\gamma_1,$$

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

$$\omega_n\left(\boldsymbol{a}^n, q_n, \mu_n\right) = \frac{\Delta_{\gamma_+}}{2\mu_n} - \frac{\Delta_{a^n}\Delta_\gamma}{2t\mu_n} q_n \tag{13}$$

$$\omega_z\left(\boldsymbol{a}^z, q_z, \mu_z\right) = \frac{\Delta_{\gamma_+}}{2\mu_z} - \frac{\Delta_{a^z}\Delta_\gamma}{2t\mu_z} q_z - \frac{\Delta_\gamma^2}{4t\mu_z^2}, \tag{14}$$

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

$$\omega_n(\mu_n) = \frac{\Delta_{\gamma+}}{2\mu_n} \tag{15}$$

$$\omega_z(\mu_z) = \frac{\Delta_{\gamma+}}{2\mu_z} - \frac{\Delta_\gamma^2}{12t\mu_z^2}. \tag{16}$$

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

$$H_n(q_n, \mu_n) = \theta q_n - \frac{3}{2}t - \omega_n(\mu_n),$$
$$H_z(q_z, \mu_z) = \theta q_z - \frac{3}{2}t - \omega_z(\mu_z) - \frac{\Delta_\gamma}{4\mu_z},$$

an thus, the ISP fully 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 the network capacity[16]. 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 parts (ii) and (ii) follows.

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

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

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

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

[16]This rent-extraction effect is at play in driving the ISPâĂŹs incentives to invest in capacity, as will be seen in the next section.

and

$$\pi_1^z = \frac{1}{2}t - \frac{(12t\mu_z - \Delta_\gamma)}{72t\mu_z^2}(\gamma_1 - \gamma_2) \text{ and } \pi_2^z = \frac{1}{2}t + \frac{(12t\mu_z - \Delta_\gamma)}{72t\mu_z^2}(\gamma_1 - \gamma_2),$$

under zero rating, where $\Delta_\gamma = \gamma_i - \gamma_j$ for $i = 1, 2$ and $i \neq j$. The contracting CP benefits from a zero rating agreement and realizes a greater level of 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 non-contracting CP. The content sector's overall profits, given by $\sum_i \pi_i^k$, are

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

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

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

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

$$\frac{\partial D_i^n}{\partial q_i^n} = \frac{\theta}{6t} \text{ and } \frac{\partial D_i^z}{\partial q_i^z} = \frac{\theta}{12t}, \tag{18}$$

and on the degree of advertising exposure

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

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, $\mathcal{BR}_i^n\left(q_j^n\right)$, 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,

$$\mathcal{BR}_i^n\left(q_j^n\right) = \frac{\theta\left(\theta q_j^n - 3t\right)}{\theta^2 - 9t} \text{ and } \mathcal{BR}_i^z\left(q_j^z\right) = \frac{\theta\mu_z\left(\theta q_j^z - 6t\right) + \theta\Delta_\gamma}{\mu_z\left(\theta^2 - 36t\right)},$$

for $i = 1, 2$ and $i \neq j$. The nature of the strategic interaction between the two CPs is then given by the signs of

$$\frac{\partial\mathcal{BR}_i^n\left(q_j^n\right)}{\partial q_j^n} = -\frac{\theta}{9t} \text{ and } \frac{\partial\mathcal{BR}_i^z\left(q_j^z\right)}{\partial q_j^z} = -\frac{\theta}{36t},$$

making clear that CPsâĂŹ qualities are strategic substitutes and that best responses are downward-sloping. Solving for both best responses and comparing incentives to invest in functionalities under each regime, we observe the following:

Lemma 3. *Suppose that $C\left(q_i^k\right) = \frac{1}{2}q_i^{k2}$. Then, for $i, j = 1, 2$ and $i \neq j$, the optimum level of 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

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

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

This lemma shows that incentives to invest in content functionalities are 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, as CPs invest the same amount in their content attributes/quality. In equilibrium, the level of vertical differentiation increases in consumers' intrinsic utility θ but is independent of the network quality provided by the ISP. However, in a discriminatory regime, investments in content quality are asymmetric, and the effect of increasing one's content functionalities is ambiguous, as it depends on the ISP's pricing strategy, and zero rating distorts CPs' market shares, creating an opportunity for them to affect demand through the quality of their content. The expressions for $q_i^{z^*}$ reflect the amplifying effect that network quality μ_z can have on the level of vertical differentiation of content. A marginal increase in network capacity increases the content quality of the non-contracting CP and decreases the contracting CP's content quality.

We now state the main result of this section. First, note that given assumption A2, we have that $\mathcal{A}\left(\theta, t, \mu_z\right) > 0$, so 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 more attractive content is, i.e., the greater the functionalities it has, the more congestion that 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 non-contracting CP reduces the ISP's ability to price discriminate between users. Indeed, the non-contracting CP makes its content more attractive by increasing its quality level, 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, investment in quality by the weak CP implies 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 congestion of its users, as it actually reduces the congestion level by doing so, but this does not sufficiently compensate for the switch. Then, the marginal effect on profits of an increase in content quality is negative for the non-contracting CP, and it opts for quality degradation because such investment is insufficient to eliminate users' switching incentives implied by the zero rating agreement.

As a result and in contradiction with 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. The investment level implied by the zero-rated CP does not sufficiently compensate for the lack of investment by its rival, so that 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 level of advertisement

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 4. *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, while the non-contracting CP reduces this level only if $\mu_z > \bar{\mu}_z(\theta, t)$.*

Let us define the network capacity threshold $\bar{\mu}_z(\theta, t) \equiv \frac{3\Delta_\gamma}{18t-\theta^2}$ and show that condition $q_i^k < \frac{3t}{\theta}$ is indeed satisfied in both regulatory regimes. From relation (19) and lemma 3, 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, $\frac{\partial a_i^n}{\partial q_i^n} < 0$. 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 4, allowing a profit-maximizing ISP to set a per-unit surcharge on users of the higher quality content implies a greater level of advertising exposure in both market segments. We obtain the following result:

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

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

and

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

In the discriminatory equilibrium with equilibrium content quality provision, advertising exposure rates are network capacity independent and symmetric across market segments. Since, in our specification, the per-unit advertising exposure rate can be interpreted as a direct payment from consumers to CPs, it also exhibits the same properties as purely paid content, where consumers pay to access the content. The result in this proposition highlights the fact that zero rating makes content more expensive to for consumers to access and reduces price competition between CPs. Consumers support greater advertisement than in a neutral network, and market shares are given by

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

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 to invest in network capacity than a neutral network, 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 and compare incentives to invest under zero rating agreements, and we find that a profit-maximizing ISP strategically reduces its network capacity through two channels.

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

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

and under zero rating agreements, they are

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

where

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

Lemma 5. *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 non-contracting CP, and (iii) increases the content quality of the non-contracting 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:

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

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

Let

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

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

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

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

$$\frac{\partial q_1^z}{\partial \mu_z} = \frac{\left(\gamma_1 - \gamma_2\right)\theta}{2\left(18t - \theta^2\right)\mu_z^2} > 0,$$

$$\frac{\partial q_2^z}{\partial \mu_z} = -\frac{\left(\gamma_1 - \gamma_2\right)\theta}{2\left(18t - \theta^2\right)\mu_z^2} < 0,$$

and the assertions in the lemma follow. □

We now derive the optimal level of network capacity under each regime and then compare their levels with the socially optimal capacity. 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

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

and

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

where

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

For the neutral network, solving for the firs-order condition of the expression for Π_n^{ISP}, we obtain that in a neutral network, the optimum network capacity 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 effect and a cost-alleviation effect.*

The proof is straightforward, as from (22) and (23), we obtain that $\left.\frac{d\Pi_n^{ISP}}{d\mu_n}\right|_{\mu_n=\mu} - \left.\frac{d\Pi_z^{ISP}}{d\mu_z}\right|_{\mu_z=\mu} = -\left.\frac{d\mathcal{S}(\theta,t,\mu)}{d\mu_z}\right|_{\mu_z=\mu} > 0$. While increasing demand for non-zero-rated content, investment in capacity actually decreases 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 content. In equilibrium, even if the overall level of quality is lower in the content market, the contracting CP is vertically more differentiated than its competitor, which makes its content more attractive. Hence, as consumers spend more time on it, this reduces the capacity of the ISP to extract rent via the per-unit fee $\tau^*(.)$ charged to consumers using the 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 ω is a negative function of the quality of the contracting content q_2. For (ii), given level of content quality $\boldsymbol{q}$ and network capacity μ, the levels of congestion are respectively given by

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

and

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

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

Next, for the first component (i), by differentiating the expressions of the optimal level of content quality derived in lemma 3 with respect to μ_z, we find that a marginal decrease in network capacity actually decreases the content quality offered by the non-contracting CP and increases the quality of the contracting CP[17]. Then, given the above 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 μ_z^*, on can observe that in response to a decrease in the non-contracting CP's quality, the profit-maximizing ISP reduces its level of investment accordingly, while the optimal network capacity is increasing in the contracting CP's quality provision[18]. However,

[17]Indeed, one obtains $\partial q_1^{z^*}/\partial\mu_z = \frac{1}{\mu_z}\mathcal{A}(\theta,t,\mu_z) > 0$ and $\partial q_2^{z^*}/\partial\mu_z = -\frac{1}{\mu_z}\mathcal{A}(\theta,t,\mu_z) < 0$.

[18]Assuming that the ISP's program is well defined and that the second-order condition is verified, one obtains that $d\mu_z^*/dq_1^z = \theta\Delta_\gamma/12t\mu_z^2 > 0$ and $d\mu_z^*/dq_2^z = -\theta\Delta_\gamma/12t\mu_z^2 < 0$.

lemma 3 indicates that zero rating implies a greater downward quality distortion from the non-contracting CP than from its competitor. As a result, the effect of this quality distortion on the level of μ_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 fails to adopt the optimal policy. We then compare the welfare-maximizing levels of 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 were to set the per-unit fee imposed by the ISP on the use of the non-zero-rated content. Consider the optimal overage fee $\tau^{\mathcal{W}}(.)$ that, given content qualities $\boldsymbol{q}_z = (q_1^z, q_2^z)$, would maximize welfare as defined by

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

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

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

leading to subgame symmetric equilibrium content qualities $q_1^z = q_2^z = \theta$. Hence, the optimal tariff that maximizes 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 τ^* greater than the welfare-maximizing two-part tariff, while if there are no congestion externalities, a profit-maximizing ISP sets the socially optimal policy, that is, $\tau^{\mathcal{W}} = \tau^* = 0$.*

This result echoes 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 0. However, in the presence of congestion, since 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 total surplus, consumer surplus and industry profits, we find that consumers are always better off under a strict regulatory regime, while total surplus is higher under net neutrality than under zero rating. In either case, we find that the content sector's profits are always higher under zero rating than under net neutrality regulation.

Proposition 6. *(i) In 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*

$\mu > 0$. *(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 $W_n(\bar{q}_n, \mu_n) > W_z(\bar{q}_z, \mu_z)$, while this relation is unclear whenever we account for equilibrium content quality provision. Computing the difference yields

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

which for given network capacity reduces to

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

This quantity is strictly positive if

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

which is always satisfied under assumption A2.

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

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

, which is positive under assumption A2.

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

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

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

While users enjoy lower access price under zero rating agreements, as $H_n > H_z$, they are subject to more advertising, i.e., $a_z^* > a_n^*$, and this supports an overall decline in content quality, as we have shown that zero rating softens CPs' quality competition in the content market, i.e., $q_i^n > q_i^z$[19]. As a result, consumer surplus is greater under net neutrality regulation. Moreover, the content industry 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. Overall, allowing zero rating agreements implies pervasive incentives

[19]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

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

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

from the ISP, as given in proposition 4, the ISP has incentives to engage in practices that make resources scarcer and always finds it profitable to financially discriminate between CPs despite that this is welfare reducing. 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, and thus welfare is lower in a discriminatory network and consumers are better off under a neutral regime.

We conclude this section by considering as corollaries the efficient levels of content qualities and of broadband provision by comparing private and social incentives to invest. We first derive the socially optimal level of content qualities and find that, contrary to private incentives, the non-contracting 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 underprovision of capacity whenever zero rating is implemented.

Efficient content quality provision. Let us compare the quality provision under both regimes $(\boldsymbol{q}_n^*, \boldsymbol{q}_z^*)$ with the socially optimal levels of quality provision $\left(\boldsymbol{q}_n^{\mathcal{W}}, \boldsymbol{q}_z^{\mathcal{W}}\right)$, which is the solution of

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

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

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

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

It is immediate that $q_i^{n^*} < q_{i,n}^{\mathcal{W}}$, so that CPs underprovide quality in the content market with respect to the socially optimal level under net neutrality. If we assume that $3t < \theta^2 < 9t$, then $\mathcal{D}(\theta, t, \mu_z) < 0$ for CP $i = 1$ and $\mathcal{D}(\theta, t, \mu_z) > 0$ for CP $i = 2$. Therefore, $q_{1,z}^{\mathcal{W}} > q_n^{\mathcal{W}} > q_{2,z}^{\mathcal{W}}$, and contrary to private incentives, the socially optimal level of quality for the non-contracting CP is greater than under net neutrality and than the zero-rated content quality. As 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 non-contracting 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 an overprovision 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, as the ISP provides capacity at the socially optimal level in the neutral network, while 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. We thus 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 $\frac{\partial \mathcal{W}_n}{\partial \mu_n} - \frac{\partial \Pi_n^{\text{ISP}}}{\partial \mu_n} = 0$, and 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 by the following corollary:

Corollary 1. *(i) CPs' investments are suboptimal, and there is underprovision in content quality. (ii) A profit-maximizing ISP's investment in broadband capacity under net neutrality is socially optimal, while (iii) the regulator has greater incentives than does the ISP to invest in broadband capacity under zero rating.*

Contrary 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 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 bilateral 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 always finds it profitable to depart from net neutrality, while zero rating is welfare reducing 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 implies a downward distortion of content quality by increasing downward vertical differentiation. Through zero rating, the ISP is able to increase the CPs' asymmetry and to affect consumers' willingness to pay for content. The CP that contracts under a zero rating plan has much greater incentives to innovate than 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 happens to reduce congestion, while 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 would provide 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 nonzero 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 would be to consider an individual congestion-quality interaction to introduce indirect network effects into the model. This would allow us to capture the idea that the more attractive content is, the more users will consume it and take time to use it and the higher the relative congestion on the network.

References

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.: 2019, Zero-rating and vertical content foreclosure, *DICE Discussion Paper* .

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)* .

Chaire Gouvernance et Régulation
Fondation Paris-Dauphine
Place du Maréchal de Lattre de Tassigny - 75016 Paris (France)
http://chairgovreg.fondation-dauphine.fr