www.erices.es

Discussion Papers
in Economic Behaviour

DPEB 24/02

Managing network traffic with discriminatory strategies: A study of zero-rating in an Internet market monopoly

> Lucía Pinar
EDEM, Spain

> Iván Arribas
Universitat de València and ERI-CES, Spain

> Amparo Urbano
Universitat de València and ERI-CES, Spain

December, 2024

Universitat de València

Managing network traffic with discriminatory strategies: A study of zero-rating in an Internet market monopoly

L. Pinar-Garcia[d], I. Arribas[a,b,c], A. Urbano[a,c,*]

[a] Universitat de Valencia, Edf. Dept. Oriental, Campus dels Tarongers, 46022 - Valencia, Spain

[b] Ivie, Calle Daniel Balaciart, 3 - bajo, 46022 - Valencia, Spain

[c] ERI-CES, Universitat de Valencia, c/Serpis 29, 46022 - Valencia, Spain

[d] EDEM, Plaça de L'aigua, Poblados Marítimos, 46024 Valencia, Spain

Abstract

Zero-rating is a recent practice that allows internet providers to discriminate among contents. We offer new insights on the role of zero-rating practices when there is network congestion. In a market where consumers differ in their preferences for network speed, those who value network speed more create a congestion externality for an Internet Service Provider that limits its ability to capture consumer surplus from those who value network speed less. The monopolist may internalize the network congestion by using two discriminatory strategies: a menu of tariffs as a screening mechanism between consumers and zero-rating as a content revaluation tool. The combination of the two strategies acts as a mechanism to control the network traffic with no incremental provision of capacity. This is done by managing the network flow, through the strategic reallocation of the network traffic between content providers and consumers. Finally, zero-rating is detrimental to consumers independently of their type.

Keywords: Internet market, Monopoly, Tariff menu, Zero-rating, Subgame perfect Nash equilibrium, Social Welfare

JEL: D43, L12, L52

⋆This work was supported by the the Spanish Ministry of Science, Innovation and Universities under project PID2019-110790RB-I00 financed by MCIN/AEI/10.13039/501100011033; and the Generalitat Valenciana under the Excellence Program Prometeo 2019/095 and CIPROM/2022/029.

*Communications to: Ivan Arribas, Dep. de Analisis Economico, Edificio departamental oriental, Avenida de los Naranjos s/n, 46022 - Valencia, Spain. e-mail: ivan.arribas@uv.es

Email addresses: lpinar@edem.es (L. Pinar-Garcia), ivan.arribas@uv.es (I. Arribas), amparo.urbano@uv.es (A. Urbano)

1. Introduction

Net neutrality is a practice by which Internet service providers ($ISPs$) should not discriminate among data packets sent on its network. Net neutrality advocates have feared that the repeal of regulation would empower Internet Service Providers ($ISPs$), which in turn would harm consumers and content providers (CPs). On the contrary, opponents to net neutrality have claimed that such an empowerment of $ISPs$ is necessary and ultimate beneficial for consumers. In recent years, two business practices, prioritization and zero-rating, have been widely discussed and debated both in policy circles and the media. Prioritization is a practice whereby an ISP creates a fast lane on the Internet network where privileged contents circulate in case of congestion, while zero-rating refers to an ISP making some content more expensive than others for consumers by price discriminating among content providers. These practices violate the net neutrality principle and allow $ISPs$ to discriminate between content in terms of quality for prioritization and in terms of price for zero-rating.

In Europe, zero-rating services have been used quite generally, with ups and downs in legislation. In 2015, the European Union (EU) decided to adopt rules like those of the US regulation, that were in place between 2015 and 2017. Specifically, BEREC (Body of European Regulator for Economic Communications) published Guidelines for net neutrality (BEREC (2016)) mandating a case-by-case treatment of zero-rating offers by national regulatory authorities (NRAs), while explicitly banning paid prioritization[1]. Recently, BEREC (2022) has banned the use of zero-rating practices. This rule has been based on the decision of the Court of Justice of EU on the cases of zero-rating offers by Vodafone and Deutsche Telekom in Germany, which violate the EU regulation on net neutrality and roaming.[2] According to a statement from the same court, the requirements of protection of the rights of Internet users and non-discriminatory treatment of traffic

[1] In practice, according to this principle, some specific zero-rating offerings were banned by NRAs (for example, the case of Telenor in Hungary and Telia in Sweden). In contrast, other offerings were found legal (Proximus in Belgium and T-Mobile in Germany).

[2] Since 2017, Telekom Deutschland offers end-customers, for some of its tariffs, an additional option (also called "add option") consisting of a free "zero-rate" tariff option called "Stream On", which allows the volume of data consumed by audio and video streaming broadcast by Telekom Deutschland's content partners not to be imputed to the volume of data included in the basic tariff. In the case of Vodafone, the zero-rating options is called Vodafone Pass, where video, music, chat and social networking applications are packaged.

are opposed to the practice of zero-rating.[3] Different reasons for the ban on zero-rating have been offered,[4] standing out among all that the possible proliferation of zero-rating offers could cause the general data limits to suffer a decrease, or, at least, discourage their increase.

Within this framework, a question that BEREC did not explicitly contemplate is the strategic use of zero-rating[5] offers by $ISPs$ to strategically manage the network traffic flow when there are transmission delays, and thus avoid network capacity expansion. To the best of our knowledge, there are not many studies that analyze this topic. Notable exceptions are Somogyi (2016), who relates congestion and increasing utility from consumption under open and exclusive zero-rating contracts, and Lorenzon (2022), who analyzes the zero-rating implications for the incentive to provide quality in the market for content and to invest in broadband infrastructure.

To study the impact of a move from the net neutrality regime on both the network flow and the capacity investment decisions of $ISPs$, we study the effect of zero-rating practices in a market where there is a monopolist ISP offering Internet services to two types of consumers, Turtle (T) and Speedy (S), the latter with a higher preference for the network flow speed than the former. There are two asymmetric content providers, CP_1 and CP_2 whose content are different in terms of quality and have a certain degree of substitutability. CPs compete in prices for their content. By considering two types of consumers' preferences for network flow speed and heterogeneous CPs, we focus on analyzing how zero-rating services can be used as a dual differentiation strategy for both Internet content providers and consumers. The ISP chooses a network capacity, and charges consumers both a connection (hookup) fee to access its network and a per unit "overage" fee for its content offerings.

The main contribution of this work to the literature is to demonstrate that in markets with consumers' heterogeneous preferences with respect to network flow speed, and with quality differentiated CPs, zero-rating combined with a tariff menu implies an increase

[3]The objectives of the Work Programme 2023 are aligned with the BEREC Strategy 2021-2025, as well as the three high-level priorities: promoting full connectivity, supporting sustainable and open digital markets, and empowering end-users.

[4]Ranging from distortion of competition between content providers, barriers to entry in content markets, the hindering innovation and development of services, etc

[5]We do not study prioritization because this practice is banned by the European Union regulation.

in content consumption without network capacity investments. This is done by managing the network flow, through the reallocation of the network traffic between content providers and consumers. The analysis allows us to precisely offer the changes in the equilibrium consumption patterns and prices paths, behind strategic zero-rating practices.

To address the analysis, our setup leans on the framework of Jeitschko et al. (2021) and of Economides and Hermalin (2015). Thus, consumers have a quadratic utility function, CPs charge consumers a per-unit fee to access their content, and demands are not unitary. In addition, we assume that both consumers and content providers are heterogeneous and there are congestion externalities. The richness of the model, combining isolated features of the received literature models, presents the disadvantage of its analytical intractability for some maximization problems. In particular, for the solution of the ISP's profit maximizing menu of tariffs. That is why we resort to computational resolutions to explore a wide range of parameter combinations and offer four types of subgame perfect equilibria of the full game that result from the strategic combination of zero-rating and tariff menus. It is important to note that our computational exercises are not mere simulations, but the computational solution of analytically intractable maximization problems, by which we are able to offer qualitative new insights.

Our first result shows that with a two-part tariff menu instead of a single tariff, the ISP designs a mechanism for traffic control, considering that a potential increase in consumption will reduce the network flow speed, which, in turn, may reduce consumption. The mechanism works as follows. On the one hand, it increases the ISP's profits by charging different pairs of hookup fee and a per unit overage fee to consumers. In this way, it charges a low per unit overage fee to the less traffic speed demanding consumers (T) and increases the fee for the more demanding ones (S). This is done by the ISP without changing the chosen network capacity with respect to the one that maximizes profits under a single two-part tariff, and which only considers the consumers with lower preferences for network speed. However, total content consumption is not increased. This is where zero-rating plays an important role, because consumption depends on content prices.

Our second result is that a departure from net neutrality combined with a tariff menu, affects even more the network traffic through additional changes in the consumption pattern. Under zero-rating, the total consumption increases with respect to net neutrality without capacity expansion. Since the less valued CP is the less demanded, by zero-rating this CP, the ISP guarantees that both types of consumers will buy from the two CPs, increasing consumption, and hence network delays. The ISP could absorb this increase in consumption with an increase in capacity. However, this is not the case, and the disutility created by delays is compensated by changes in the consumption pattern, which take advantage of the new CPs' relative prices. Namely, in the new consumption pattern, the less speed demanding consumers may buy more from the more attractive CP, and less from less valued one (as being like an inferior good for these consumers), and the more speed demanding consumers buy more from the less valued CP.

The third and completely novel result shows that the subgame perfect Nash equilibria with zero-rating are separated from those without zero-rating by a frontier, which depends mostly on the degree of substitution between contents and the proportion of consumers of each type. In addition, we find that the greater the substitution between content, the greater the heterogeneity between content providers, and the lower the consumers' differentiation, the greater the range of parameter profiles that leads to the existence of zero-rating equilibria.

Finally, the social welfare analysis shows that zero-rating the less valued CP by the ISP generates greater social welfare than no zero-rating at all, unless CPs are highly substitutable. However, consumers and the no zero-rated CP are worse off under zero-rating. Hence, a profit-maximizing ISP has pervasive incentives to implement zero-rating contracts, and these incentives are misaligned with consumer welfare.

Hence, our results support BEREC's decision to ban zero-rating practices in the EU. With such a decision, the regulatory body prevents traffic distorting practices that harm the end user as well as anti-competitive effects on smaller content providers.[6]

[6] Article 3(3) third subparagraph of the Open Internet Regulation (BEREC (2022)) sets out traffic management practices that are banned, and can be described by seven basic principles, which should be used by NRAs when assessing ISPs' practices. Between specific content, applications or services, or specific categories thereof, there should be: no blocking; no slowing down; no alteration; no restriction; no interference with; no degradation; and no discrimination.

1.1. Related literature

Economic studies on zero-rating practices have been appearing in recent years. Much of the literature considers that CPs obtain revenues from advertising and passively competing to attract consumers. In Inceoglu and Liu (2019), zero-rating is implemented to screen among consumers in an environment with multiproduct demand. Zero-rating is found to be welfare-enhancing and causes network capacity expansion. Jullien and Sand-Zantman (2018) consider zero-rating contracts as a coupon from content providers to end-users, a potential way to overcome the misallocation problem stemming from the free content business model. Here, we analyze the combination of zero-rating services with different pricing schemes: either single two-part tariff or a menu of two-part tariffs. In the context of sponsored data, Jullien and Sand-Zantman (2018) argue that the price structure is irrelevant for the market outcome. However, we demonstrate the practical relevance of zero-rating for paid content as a pricing strategy.

Several authors have studied the effect of zero-rating in models where CPs do not set a price for their content. Here, we do not consider payments from CPs to $ISPs$, and CPs set prices for their content consumed by consumers. In Gautier and Somogyi (2020) zero-rating is viewed as a tool to enhance content differentiation and to alter the competition in the market for content. However, their 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. Our aim is to study the implications that a departure from a network neutrality regulation has for $ISPs$' incentives to provide broadband facilities and to redirect the network traffic flow. Most of the studies addresses the effect of zero-rating services on social welfare. Preta and Peng (2020) find that while zero-rating could have anti-competitive consequences, it is potentially welfare enhancing based on network effects, price discrimination, two-sided market and behavioural economics explanations. Along with this, Eisenach (2015) also describes the economic foundations of zero-rating and its concerns, and concludes that zero-rating is a mechanism to capture economic efficiencies, while rejecting it as an anti-competitive strategy. Inceoglu and Liu (2019) focus on the consumer side of the Internet market and model zero-rating as a screening device for price discrimination. They found that generally, zero-rating can increase welfare if contents are not close substitutes and preferences are sufficiently different. We also find that zero-rating increases welfare when content are not close substitutes, but consumers and the no zero-rated CP are worse off.

Like us, Schnurr and Wiewiorra (2018) also study the effects of zero-rating services in a monopolistic market with heterogeneous consumers. They analyze the zero-rating offers and sponsored data based on firms' profits and prices, as well as consumer surplus and total welfare. Schnurr and Wiewiorra (2018) find that the combination of data caps with zero-rating can be used to discriminate between different types of consumers depending on the level of heterogeneity. Although we do not consider data caps, we approach them in the sense of finding that the use of a tariff menu and zero-rating services can be used as a double discrimination tool. Finally, the zero-rating mechanism as a price discrimination tool is also related to the literature on the strategic use of capacity limits (Hoernig and Monteiro, 2020). More generally, our work contributes to the understanding of multi-product price discrimination (Armstrong, 1996), albeit in a specific context.

The two papers relating congestion and zero-rating are Somogyi (2016), who relates congestion with open and exclusive zero-rating contracts and Lorenzon (2022) who analyzes the zero-rating implications for the incentive to provide quality in the market for content and to invest in capacity. Lorenzon (2022) considers a unit mass of consumers with unit demand, and CPs competing à la Hotelling to offer vertically differentiated services to consumers under an advertising-supported service model. On the contrary, as in Jullien and Sand-Zantman (2018) and Jeitschko et al. (2021), we examine paid content, and, as in Schnurr and Wiewiorra (2018), heterogeneous consumers.

In what follows, we set up our model in Section 2. The equilibria of the baseline model under a tariff menu, where the monopolist does not offer zero-rating services is derived in Section 3. In Section 4 our monopolist is allowed to follow a zero-rating strategy. In Section 5 we offer the subgame perfect equilibria of the game, while Section 6 analyzes the equilibrium consumption patterns and network capacity. Finally, Section 7 analyzes the welfare implications and Section 8 concludes the study.

2. The Model

The model consists of a single Internet service provider (ISP), two content providers (CP_1 and CP_2) differentiated by quality, and a unit mass of two types of consumer, Turtle and Speedy (T-consumer and S-consumer, respectively) who differ from each other in their sensitivity to the network connection speed. The proportion of T-consumers in the market is $\theta \in (0,\ 1)$.[7]

We consider two network regimes: the neutral network, where Internet data packages are treated equally, and the restricted network where the ISP establishes an exclusive zero-rating contract with one of the CPs.

Internet service provider (ISP)

Under both regimes, the ISP chooses a network's capacity μ or network bandwidth[8] at a cost. Let $C(\mu)$ be the cost of building a network capacity μ, with $C(\mu) = c \cdot (\mu + \frac{1}{2}\mu^2)$ where c, the parameter associated with the network capacity, is small enough to guarantee that the ISP can build up enough capacity to serve both consumers' type. The ISP also charges a two-part tariff to the m-consumer ($m = T, S$): a hookup fee, denoted H^m, and a linear per unit "overage" charge, denoted τ^m, for their consumption content. Both the hookup fee and the overage fee could be the same for the two types of consumers or could be different for each type of consumer. As in Jeitschko et al. (2021), Lorenzon (2022) and other authors we are interested in situations where consumers exceed the cap, therefore we assume that the data cap is set to zero. Under the restricted network regime, the ISP has the option to offer zero-rating to one of the CPs. If the CP accepts the zero-rating service, the consumption of such zero-rated content will not be counted in the variable part of the two-part tariff paid for Internet access by both type of consumers. The ISP

[7]This proportion will capture the different segments of Internet users in the population with different network speed needs.

[8]Network bandwidth is a measurement that indicates the maximum capacity of a communication link to transmit data over a network connection in a given amount of time and is typically represented by the number of bits that can be transmitted in 1 second. Synonymous with capacity, bandwidth describes the speed of data transfer. Bandwidth is not a measure of network speed, a concept with which it is often confused.

does not charge any fee to the CPs. The ISP's profit function is specified by,

$$\pi^{ISP} = \begin{cases} \theta[H^T + \tau^T(q_1^T + q_2^T)] + (1-\theta)[H^S + \tau^S(q_1^S + q_2^S)] - C(\mu) & \text{if neutral network,} \\ \theta[H^T + \tau^T q_i^T] + (1-\theta)[H^S + \tau^S q_i^S] - C(\mu) & \text{if zero-rating with } CP_j. \end{cases} \tag{1}$$

where q_j^m is the m-consumer's usage volume corresponding to content provided by CP_j, $j = 1,\ 2$ and $m = T,\ S$.

Content providers (CPs)

The two CPs offer different content quality and bear no cost associated to content production. The content quality of CP_j, $j = 1, 2$ is measured by parameter α_j, with $1 = \alpha_2 < \alpha_1 \leq 2$. This condition implies that the quality of CP_1's content is up to twice higher than that of the lower quality CP_2. In addition, the content of each CP is either a substitute for or independent of that of the other one, with $\gamma \in (0, 1)$ denoting the degree of content substitutability. Thus, this parameter measures the closeness of competition between the two CP's content.[9] A high value of γ indicates that each CP content is a relatively close substitute for that of the other one, whereas a low value indicates the opposite. In what follows, we restrict the degree of content substitutability, γ, to guarantee that both CPs have positive market shares in equilibrium (i.e., interior solutions). This requires γ to be sufficiently low so that both consumer types buy from at least one CP. Thus, each CP content is generally not perfectly substitutable by that of the other CP. In addition, we assume that CPs are insensitive to network speed.

Under both regimes, the CPs set simultaneously and independently usage fees, p_j, $j = 1, 2$, for their content to consumers. Under the restricted network regime, each CP_j must decide first whether to accept a zero-rating offer, if any. We assume that CPs are not sensitive to the network speed. Then, each CP_j profit function is defined by,

$$\pi_j^{CP} = p_j[\theta q_j^T + (1-\theta) q_j^S]. \tag{2}$$

[9]For instance, it captures that cloud service providers like Google Cloud and Amazon web services are perceived as competitors, meanwhile a music streaming service and a film streaming services are more differentiated.

Consumers

Consumers obtain utility from their content consumption. Due to some technological constraint on capacity, traffic from the content providers to the end users might suffer from some delay as the content consumption increases. As in Bourreau et al. (2015) and Choi and Kim (2010), this delay is measured by the waiting time[10] for end users when they request content from CPs. As common in the literature, we adopt the M/M/1 queuing model to determine the average level of delay as a function of network capacity and traffic.[11] Then, the network expected waiting time (i.e., the average level of congestion) is

$$w = \frac{1}{\mu - [\theta(q_1^T + q_2^T) + (1-\theta)(q_1^S + q_2^S)]}$$

where, to have w well defined, $\mu - [\theta(q_1^T + q_2^T) + (1-\theta)(q_1^S + q_2^S)] > 0$. Thus, the expected waiting time decreases with the network capacity while it increases with the consumption of content in the network.

The net utility function U^m of the consumer of type $m = T, S$ is defined by a variation of the quasi-linear quadratic utility[12] function:

$$U^m = u(q_1^m, q_2^m) + \frac{d^m}{w} - \left(p_1 q_1^m + p_2 q_2^m\right) - HT^m, \tag{3}$$

where

$$\begin{aligned} u(q_1^m, q_2^m) &= \alpha_1 q_1^m + \alpha_2 q_2^m - \frac{1}{2}[(q_1^m)^2 + (q_2^m)^2 + 2\gamma q_1^m q_2^m], \\ HT^m &= \begin{cases} H^m + \tau^m(q_1^m + q_2^m) & \text{if neutral network,} \\ H^m + \tau^m q_i^m & \text{if zero-rating with } CP_j, \end{cases} \end{aligned}$$

and $d^m \in (0,\ 1)$ is a parameter measuring the m-type consumers' sensitivity to the

[10]The concept of waiting time is similar to that of latency, which is a more appropriate term for networks to describe the total time it takes for a data packet (information) to travel from one source node to another destination node. Latency is usually measured in milliseconds (ms). Thus, the lower the number of milliseconds, the more efficient the network behavior.

[11]On the M/M/1 model, see Choi and Kim (2010) and the references cited therein.

[12]See Choné and Linnemer (2020).

network speed connection. We assume that $d^T < d^S$, that is, S-consumers are more sensitive to the network speed than T-consumers.[13] Thus, as in Bourreau et al. (2015), $\frac{d^m}{w}$ represents the preference (utility) for the connection speed, that is, for the network flow speed, which by above can be expressed as $d^m[\mu - \theta(q_1^T + q_2^T) - (1-\theta)(q_1^S + q_2^S)]$.

S-consumers, who value network speed highly, create a negative congestion externality for the ISP that limits its possibilities for capturing consumer surplus from those who value the network speed less. The monopolist will also try to internalize the negative congestion externality by either building a high capacity μ or by inducing S-consumers to buy less content, by charging different consumption prices. The latter can be done by either designing a two-part tariff menu or by zero-rating offers with or without a tariff menu.

Timing

The model is a three-period game with complete information: In the first period, the ISP must decide whether to offer zero-rating to one of the CPs in a first stage, then it chooses a single network capacity μ and offers to consumers a menu of two-part tariffs $(\{H^T, \tau^T\}, \{H^S, \tau^S\})$, which can be a single two-part tariff. In the second period, each CP decides in a first stage whether to accept the zero-rating offer, if any, then both CPs choose simultaneously and independently prices p_j for their contents, $j = 1,\ 2$. In the third period, the two types of consumers decide the content quantities to purchase q_j^m, $j = 1,\ 2$ and $m = T,\ S$.

We study the subgame perfect equilibria of this sequential game under the two regimes: the neutral network (no zero rating) and the restricted network (zero-rating).

3. Neutral network: No zero-rating

First, we derive the equilibria of a neutral network, i.e., the ISP does not offer zero-rating to any CP. As usual we solve by backward induction.

[13]Note that consumers' sensitivity to the network speed connection can be interpreted as the inverse of latency. For example, online gaming applications require a low latency, typically less than 50 ms, while video conferencing applications can tolerate higher latency, up to 150 ms. The former (gamers) will belong to the segment of users with a d^S parameter, while the latter (users of Zoom, small firms) will be of the d^T type.

3.1. Third period: Consumers' equilibrium strategies

The $m-$consumer's equilibrium strategies are the content quantity of no-negative consumption q_1^m and q_2^m. Consumers are aware of congestion and each type of consumer internalizes the congestion of their own group. Then, each consumer maximizes her utility function, given the ISP and the CP's equilibrium pricing strategies:

$$\begin{aligned} Max_{q_1^m, q_2^m} \quad & U^m(q_1^m, q_2^m | p_1, p_2, \mu, H^T, H^S, \tau^T, \tau^S) \\ s.t. \quad & q_1^m, q_2^m \geq 0 \end{aligned}$$

The first order conditions give us the equilibrium consumption functions that each type of consumer makes of each of the content providers, for $i, j = 1, 2,\ i \neq j$. The second order condition implies that $\gamma < 1$.

$$q_i^T = \frac{1}{1-\gamma^2}\Big[\alpha_i - \gamma\alpha_j - (p_i - \gamma p_j) - (1-\gamma)[\theta d^T + \tau^T]\Big] \tag{4}$$

$$q_i^S = \frac{1}{1-\gamma^2}\Big[\alpha_i - \gamma\alpha_j - (p_i - \gamma p_j) - (1-\gamma)[(1-\theta) d^S + \tau^S]\Big] \tag{5}$$

The T-consumers' utility function depends on that of S-consumers through the network flow speed term and vice versa. However, due to the linearity of demands, the T-consumers' demand function does not depend on the S-consumers' content consumption. The same is true for S-consumers. Therefore, the congestion externality only operates on the utility function through the network flow $(\mu - \theta(q_1^T + q_2^T) - (1-\theta)(q_1^S + q_2^S))$, or the average level of congestion w. The m-consumers' demand function depends on CP's equilibrium prices, their sensitivity to the network speed flow d^m, and the ISP's unit overage charge τ^m. Also q_i^T decreases with θ, meanwhile q_i^S increases with it.

Let θd^T be the *marginal delay cost* of T-consumers' consumption and $(1-\theta)d^S$ be that of S-consumers. Moreover, let $\theta d^T + \tau^T$ be the *marginal Internet usage cost* of T-consumers and $(1-\theta)d^S + \tau^S$ be that of S-consumers. From equations (4) and (5), $q_j^T - q_j^S = \frac{1}{1+\gamma}[(1-\theta)d^S + \tau^S - (\theta d^T + \tau^T)]$. Thus, the relationship between consumption depends on the marginal Internet usage costs.

Lemma 1. *The T-consumers' demand from each CP_j will be higher than that of the S-consumers if and only if the marginal Internet usage cost of T-consumers is lower than the marginal Internet usage cost of S-consumers, i.e.,* $\theta d^T + \tau^T < (1-\theta)d^S + \tau^S$.

Therefore, the T-consumers' demand of each CP's content will be higher than that of S-consumers, whenever $(1-\theta)d^S > \theta d^T$, given that $\tau^T < \tau^S$, as we will see later on.

3.2. Second period: Content Providers' equilibrium strategies

The content providers' equilibrium strategies are to set simultaneously and independently the prices that maximize their profit functions, given the ISP equilibrium strategy, and consumers' equilibrium quantities in the continuation game, and for i, $j = 1, 2$:

$$\begin{aligned} Max_{p_j} \pi_j^{CP}(p_j \mid p_i, \tau^T, \tau^S) &= p_j[\theta q_j^T + (1-\theta) q_j^S] \\ s.t. \qquad & p_j > 0. \end{aligned}$$

By subgame perfection, substituting q_j^T and q_j^S by equations (4) and (5) (the demand functions of period 3), and solving the system,

$$p_1 = \frac{1}{4-\gamma^2}\left[(2-\gamma^2)\alpha_1 - \gamma\alpha_2 - (1-\gamma)(2+\gamma)(\bar{d}+\bar{\tau})\right] \tag{6}$$

$$p_2 = \frac{1}{4-\gamma^2}\left[(2-\gamma^2)\alpha_2 - \gamma\alpha_1 - (1-\gamma)(2+\gamma)(\bar{d}+\bar{\tau})\right] \tag{7}$$

where $\bar{\tau} = \theta\tau^T + (1-\theta)\tau^S$ is the average per unit overage charge, and $\bar{d} = \theta^2 d^T + (1-\theta)^2 d^S$ is the average marginal delay cost. The CPs will consider the average of both consumer types' marginal delay cost when setting prices, which will be lower than those without consumers' preferences for the network speed flow ($d^T = d^S = 0$). These lower prices seek to increase consumers' demand, compensating them for the waiting time. Moreover, $\bar{d} + \bar{\tau}$ is the *average marginal Internet usage cost*, and the higher this usage cost the lower CPs' prices.

Some straightforward calculations show that $p_1 - p_2 = \frac{1+\gamma}{2+\gamma}(\alpha_1 - \alpha_2)$ which implies that $p_1 > p_2$. The economic intuition is clear, the higher quality CP can afford to set the higher price, and the higher the difference in quality the greater the difference between prices. Furthermore, the higher the content substitution the higher this difference, since the lowest quality CP needs to set a lower price in an attempt to capture consumers.

In addition, $q_1^m - q_2^m = \frac{1}{(1-\gamma)(2+\gamma)}(\alpha_1 - \alpha_2)$ which implies that $q_1^m > q_2^m$. At equilibrium, any type of consumer demands more from the high quality CP, and the higher the difference in content quality the greater the difference in equilibrium consumption. Also, the higher the degree of substitution between content the greater the consumers' preference for the high quality content.

Note that, if there were no quality differentiation of CPs (i.e., $\alpha_1 = \alpha_2$), then CPs prices would be the same and each consumer's type would consume the same from the two CPs.

3.3. First period: Internet Service Provider's equilibrium strategies

In the first period, the ISP builds the network capacity, μ, and may discriminate between both consumer types via self-selection. Specifically, the ISP offers a portfolio of two-part tariffs, $(\{H^T, \tau^T\}, \{H^S, \tau^S\})$. Let us briefly explain the logic of the argument behind the ISP's discrimination policy. Consumers with a high network speed value (S-consumers) create a congestion externality for the monopolist that limits the possibility for capturing consumer surplus from those who value less the network speed (T-consumers). The monopolist internalizes this effect by inducing S-consumers to consume smaller amounts of both content, opening the possibility of charging lower prices to T-consumers to boost their consumption.

The ISP's equilibrium strategies are obtained by solving the following optimization problem, where the ISP maximizes its profit subject to the incentive and participation constraints,

$$
\begin{aligned}
Max_{H^T,H^S,\tau^T,\tau^S,\mu}\ \pi^{ISP} &= \theta[H^T+\tau^T(q_1^T+q_2^T)]+(1-\theta)[H^S+\tau^S(q_1^S+q_2^S)]- \\
&\quad c(\mu+\frac{1}{2}\mu^2)
\end{aligned}
$$

$$
\begin{aligned}
s.t. \quad & U^T(q_1^T(\tau^T),q_2^T(\tau^T)) \geq U^T(q_1^T(\tau^S),q_2^T(\tau^S)) && (8)\\
& U^S(q_1^S(\tau^S),q_2^S(\tau^S)) \geq U^S(q_1^S(\tau^T),q_2^S(\tau^T)) && (9)\\
& U^T(q_1^T(\tau^T),q_2^T(\tau^T)) \geq 0 && (10)\\
& U^S(q_1^S(\tau^S),q_2^S(\tau^S)) \geq 0 && (11)\\
& H^T,H^S,\tau^T,\tau^S,\mu \geq 0,
\end{aligned}
$$

where we have assumed that the outside option for each consumer's type is normalized to be zero. Incentive constraints (8) and (9) guarantee that the consumption by any consumer type is not distorted. That is, the menu of two-part tariff designed for T-consumers gives those consumers a utility greater or equal than the menu designed for S-consumers; and similarly, the two-part tariff menu designed for S-consumers gives those consumers a utility greater or equal than that designed for T-consumers. On the other hand, participation constraints (10) and (11) guarantee that both consumer types subscribe with the ISP.

We start by analytically solving the optimal value of ISP's decision variable μ, the network capacity. As in standard mechanism design problems, assume that the S-consumers' incentive compatibility constraint (the high type) and the T-consumers' participation constraint (the low type) are binding, then

Proposition 1. *If only constraints (9) and (10) are binding, then* $\mu = d^T/c - 1$.

Proof. If only constraints (9) and (10) are binding, then, we will proceed as follows: First, from constraint (10) we can solve for decision variable H^T as a function of the other decision variables and market parameters and, in the same way, from constraint (9) we can solve for decision variable H^S. Second, we substitute H^T and H^S in the objective function by the expression obtained in the first step. Some algebra shows that the objective function takes the form $\pi^{ISP} = g(\ldots) + d^T\mu - c(\mu + 0.5 \cdot \mu^2)$, where the function g does not depend on μ. Thus, the optimal value of μ is independent of the other decision variables. Finally, we maximize the objective function with respect μ to get its optimal value $\mu = d^T/c - 1$.[14] □

The above solution would coincide with the standard result of marginal cost equating marginal revenue in Gautier and Somogyi (2020), if the authors adopted our network cost function. Thus, the optimal network capacity is independent of CPs' characteristics: Degree of quality (α_1 and α_2) and degree of substitutability (γ); and it is also independent of the distribution of consumers' type (θ). The optimal μ only depends on the ISP's cost associated to the network capacity (c) and the marginal revenue of the lowest consumers' type preference for network speed (d^T). Therefore, the optimal investment in network capacity becomes a long-term decision.

Now, we turn to solving the two-part tariff menu. Both the objective function and the constraints of the above maximization problem are quadratic functions of the decision variables τ^T and τ^S. This makes the problem intractable so that it is impossible to derive algebraically a closed-form solution for the ISP decision variables H^T, H^S, τ^T, and τ^S. This is still true even if we assign precise numerical values to the parameters of the problem.

For this reason computational methods are essential for solving the ISP maximization problem.[15] Therefore, we examine a broad range of computational instances of the above maximization problem in which we solve for the optimal two-part tariff menu for different sets of the market parameters $\alpha_1, \alpha_2 = 1, \gamma, d^T, d^S, \theta$ and c. To enable reproducibility of our results, in Appendix A we define the grid of uniformly spaced parameter combinations used.

[14] The above Proposition is valid to the extent that only constraints (9) and (10) are binding. This happens for a wide range of the market parameter values. However, for some few configurations also constraint (11) is binding. This situation may take place under a low market share of T-consumers and low-capacity costs. In this case, the optimal value for μ would become more complex, and depend on the other ISP decision variables, and ultimately on all market parameters. Nevertheless, it would be, in general, higher than $d^T/c - 1$.

[15] As mentioned in the introduction, this methodology is different from mere simulations. It is about using the computational methodology to solve a maximization problem that is analytically intractable, but which allows us to give new insights into more complex real-life problems.

The methodology is as follows. Given a set of market parameters, and after solving the maximization problem, we obtain the optimal per unit linear overage charge tariffs τ^T and τ^S. Then, these values are plugged into the equilibrium consumption functions (4)-(5) and the equilibrium price functions (6)-(7) to get their maximizing values. Now, we check whether the set of market parameters lead to both *interior* and *valid* solutions. A set of market parameters leads to *interior* solutions if: i) The two equilibrium prices are positive; ii) all demands are non-negative; and iii) both CPs sell content in equilibrium. Thus, in the optimal ISP's strategy both types of consumers are active in the market.[16] A set of market parameters leads to *valid* solutions if: i) Consumers obtain non-negative utility; and ii) the network has a non-negative level of congestion, i.e., $\mu \geq \theta(q_1^T + q_2^T) - (1-\theta)(q_1^S + q_2^S)]$. All these requirements impose an upper bound in the parameters. For example, if c were big, then the optimal hookup fee would have to be set so high than consumers' utility would be negative; if either d^T or d^S were big enough, then the level of congestion would always be negative. Thus, we have to set an upper bound for parameters c, d^T and d^S so that no interior and valid solutions are excluded from the computational analysis, but outside the range of the analyzed set of marker parameters there are no interior and valid solutions.

Fortunately, both the objective function and the constraints exhibit a smooth behavior with respect to the parameters and the decision variables (i.e., the objective function and the constraints are polynomial, and the binding constraints are the same for all instances of the problem) which translates to the solutions of the maximization problem. In other words, the optimal solutions of the maximization problem will exhibit a continuous and smooth behavior with respect to model parameters.[17]

It is important to acknowledge the limitation of this computational analysis. Although we have considered a wide parameter space, the results will clearly depend on our final

[16]It could be the case that the ISP maximizes its profit by allowing only one type of consumer to be active. These new scenarios will increase the computational analysis to such an extent that will be beyond the scope of this paper. Moreover, in real world it is observed that all interested consumers can access the Interner content.

[17]The maximization problems have been solved using function *nloptr* from the R package of the same name. R is a language and environment for statistical computing, and the function nloptr is the R Interface of NLopt. NLopt is a free/open-source library for nonlinear optimization, providing a common interface for a number of different free optimization routines available online as well as original implementations of various other algorithms. See https://nlopt.readthedocs.io/en/latest/NLopt_Algorithms for more information.

choice. Moreover, we cannot assure that the chosen parameter combinations cover all relevant solutions of interest, although the extensive grid of parameters combinations and the smooth behaviour of the functions involved give us some guarantee that this is the case. Even so, the comparative analysis of the solutions will allow us to draw conclusions regarding the equilibrium strategies followed by consumers, content providers and the Internet service provider. In fact, we will see that our solutions coincide with the algebraic ones in tractable mechanism design problems. The following Result states the solution of the two-part tariff menu discrimination problem. [18]

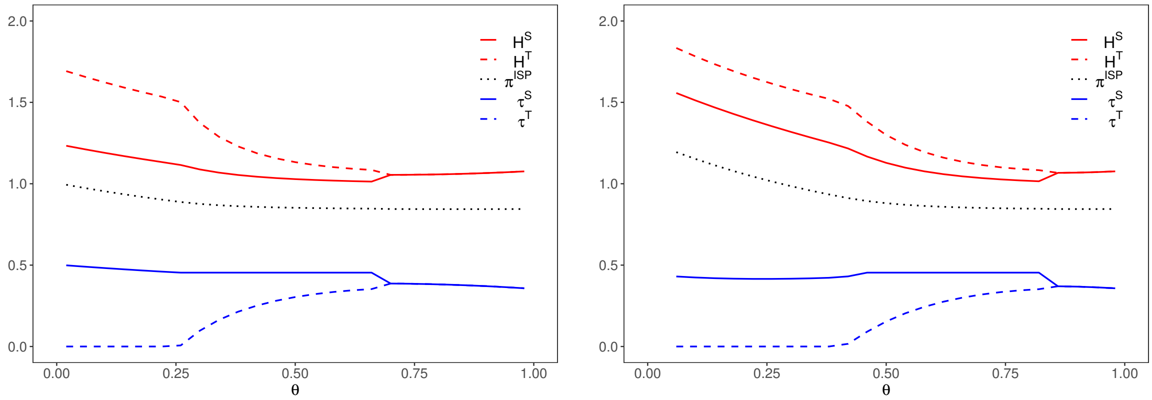
Result 1. *At equilibrium, the ISP designs an optimal interior two-part tariff menu such that: i) the T-consumers' surplus is fully extracted (constraint (10) is binding); and ii) S-consumers are indifferent between choosing any of the two-part tariff menu (constraint (9) is binding).*

The computational analysis of the equilibrium two-part tariff menu reveals two main results: First, under a menu of tariffs, the ISP always offers to S-consumers a tariff with a positive hookup fee and a per unit overage charge, with the former being lower than that for T-consumers ($H^S < H^T$) and the latter being higher ($\tau^S > \tau^T$). These results are the usual ones in mechanism design. Second, the ISP does not need to always offer a menu of tariffs to extract the maximum profit and it could be optimal to set the same tariff to both consumer types, so that in the equilibrium strategy $H^T = H^S$ and $\tau^T = \tau^S$.

The analysis shows that the relationship between the marginal delay costs $(1-\theta)d^S$ and θd^T is the relevant argument for the ISP's equilibrium strategy. Let us illustrate these results in Figure 1, which shows the equilibrium ISP's hookup fees, the per unit overage charges and profits depending on the distribution of consumer types and the degree of consumers' differentiation d^T and d^S. The left hand side of the figure refers to the case with almost homogeneous consumers (i.e., $d^T = 0.1$ and $d^S = 0.2$) and the right hand side of the figure to that of high consumer differentiation ($d^T = 0.1$ and $d^S = 0.5$). Both figures reveal a similar pattern: When the share of T-consumers is below a certain proportion, the ISP's equilibrium strategy is to offer a different two-part tariff to each type of consumer; and above that proportion the ISP offers both consumer the same

[18] We name as Lemma and Proposition the algebraic results, and as Result those coming from computational solutions.

Figure 1: ISP' equilibrium two-part tariff menu under no zero-rating



(a) $d^S = 0.2$ (Low consumer differentiation)

(b) $d^S = 0.5$ (High consumer differentiation)

Notes: Two-part tariff menu has been computed for parameters $\alpha_1 = 1.5$, $\alpha_2 = 1$, $\gamma = 0.2$, $d^T = 0.1$ and $c = 0.01$, and scale has been set equal for both plots.

two-part tariff.

The economic intuition is clear. As long as the ISP does not extract all the S-consumers' surplus, it builds an equilibrium network capacity according to T-consumers' preferences for network speed. Now, consider a τ, common to both consumers' types. The extent of the negative congestion externality in consumption of S-types is given by $(1-\theta)d^S$. Therefore, if $(1-\theta)d^S > \theta d^T$, then it is more costly in terms of network delays an increase in the demand of S-consumers than an increase of that of T-consumers, and by lemma 1, T-consumers will demand more from both CPs than S-consumers. To internalize the S-consumers congestion externality on consumption, the ISP helps T-consumers to consume more, by charging them a low τ^T and extracting all their surplus through a high H^T. In fact, for very low T-consumers' cost of waiting time θd^T, the ISP could charge a $\tau^T = 0$ to boost their demands. On the contrary, and to counter this strategy, the ISP charges S-consumers a higher τ^S, but a lower H^S, to fulfill their incentive constraint.

Eventually, beyond the T-consumers' proportion that makes $(1-\theta)d^S = \theta d^T$, i.e. $\theta = d^S/(d^S + d^T)$, the ISP's equilibrium strategy is to offer both consumer types the same two-part tariff. This threshold increases as the consumers' differentiation increases (in the left hand side of Figure 1, with $d^T = 0.1$ and $d^S = 0.2$, it happens when $\theta = 0.67$, and in the right hand side where $d^S = 0.5$ when $\theta = 0.83$). Offering a single two-part tariff now is optimal due to the fact that the ISP's profit decreases as the share of T-consumers increases (see black dashed line in Figure 1), because of the decrease in both the T-consumers' consumption and the hookup fee, which are never compensated by the increase of the T-consumers' per unit overage charge. Moreover, as Lemma 1 states, when the proportion of T-consumer is high enough and then $(1-\theta)d^S < \theta d^T$, they consume less from any CP than S-consumers. Thus, it is very difficult for the ISP, to profitably internalize the congestion disutility of S-consumers and hence to separate them from T-consumers. The ISP gains more by bunching them onto a single two-part tariff. The specific details of consumers' demands will be seen in Section 6 (Equilibrium consumption patterns and network capacity).

In addition, as already mentioned, if there were no quality differentiation of CPs (i.e., $\alpha_1 = \alpha_2$) then, by equations (6) and (7), CPs prices would be the same and, by equations (4) and (5), each consumer's type would consume the same from the two CPs. In that case, it would be very costly for the ISP to manipulate the consumption pattern and then, it would set a single two-part tariff at equilibrium. Thus, for a very low CPs' quality differentiation, a single two-part tariff will allow the ISP to extract all the T-consumers' surplus. On the contrary, for high CPs' differentiation, and heterogeneous consumers a tariff menu may allow the ISP to profitably separate consumers even more.

A detailed discussion of the role of the discriminatory tariff with respect to the ISP's investment in network capacity and to the network traffic flow compared to the single two-part tariff is offered in detail in Appendix B.2. A first result is that in both tariff schemes the ISP chooses the same optimal network capacity, which is sensible given that the marginal costs of building capacity and the marginal revenues are the same. Therefore, the ISP may redirect the single two-part tariff traffic flow in its own benefit to deal with S-consumers' congestion externality.[19]

To sum up, to cope with congestion externalities the ISP operates through network traffic redirection by designing a two-part tariff menu instead of a single two-part tariff. To do that, the ISP forces changes in the content's relative prices to manage traffic flow. These relative prices are the sum of τ's and CP's prices. However, the design of the different τ's by the ISP has to fulfil the incentive compatibility constraints, which impose some bounds to its optimal choice. Namely, for any given pair d^T and d^S, there is a proportion of T-consumers (θ^*) such that the average optimal overage fee, $\bar{\tau} = \theta^* \tau^T +$

[19]Note that since the network capacity is the same this is a short-term analysis as well.

$(1-\theta^*)\tau^S$ is equal to the optimal single two-part tariff overage fee, and then content providers' prices are the same under the two tariff schemes (see equations (6), (7), and Appendix B.1). Therefore, from θ^* on there are not changes in the CP's relative prices which would help to redirect content consumption between the two types of consumers, and only the design of the τ's must make the job. But this design may also face difficulties. Specifically, from a $\hat{\theta} = d^S/(d^S + d^T)$ on, where ($\hat{\theta} > \theta^*$), the content consumption of the two type of consumers is the same in both schemes, and then the two-part tariff menu is not useful at all. The ISP can only choose a single two-part-tariff.

Therefore, given the above economic intuition, the ISP could appeal, when profitable, to zero-rating offers to change the relative CP's prices and assist the ISP in the design of the optimal τ's and, hence, in the network flow redirection. Thus, two-part tariff menus and zero-rating (when profitable) are complementary strategies to internalize network congestion and, in particular, the S-consumer congestion externality. Nevertheless, even with zero-rating, the two-part tariff menu also has constraints to fulfil and this is why zero-rating offers may co-exist with a single two-part tariff scheme.

4. Restricted network: Zero-rating services

In this Section, we study the subgame perfect equilibria of the sequential game when, in the first stage of the first period, there is an exclusive zero-rating contract between the ISP and one of the CPs. For ease of presentation, we will only explain the model when the ISP zero-rates CP_1. However, all results in this section are offered for both cases: when the ISP zero-rates CP_1 and when ISP zero-rates CP_2. Algebraic calculations are in Appendix C.1 and Appendix C.2.

As the ISP zero-rates the content of CP_1, this content is not part of any consumers' tariff. The m-consumers pay a fixed part, H^m, and the variable part linked to their CP_2's content consumption. We derive the subgame perfect equilibrium by backward induction.

Third period: The consumers' equilibrium strategies.

The m-consumer net utility function is given in (3) where $HT = H^m + \tau^m q_2^m$. Then, each consumer's type maximizes her own net utility function, subject to non-negative quantities and to her income constraint, and given the ISP and CP's equilibrium strategies.

Under zero-rating, the relationship between T-consumers' equilibrium consumption of any CP and that of the S-consumers becomes more complex. From (C.1) to (C.4) in Appendix C, we have that,

$$q_1^T - q_1^S = \frac{1}{1-\gamma^2}\Big[(1-\gamma)[(1-\theta)d^S - \theta d^T] - \gamma(\tau^S - \tau^T)\Big] \quad (12)$$
$$q_2^T - q_2^S = \frac{1}{1-\gamma^2}\Big[(1-\gamma)[(1-\theta)d^S - \theta d^T] + (\tau^S - \tau^T)\Big] \quad (13)$$

As we will see later on, at equilibrium $\tau^S > \tau^T$. This fact allows us to prove that if $(1-\theta)d^S < \theta d^T$, then $q_1^T < q_1^S$; and that if $(1-\theta)d^S > \theta d^T$, then $q_2^T > q_2^S$. In general, if the proportion of T-consumers is above the threshold $d^S/(d^S + d^T)$, then T-consumers' consumption of the zero-rated CP will be lower than that of S-consumers. On the contrary, if the proportion of T-consumers is below this threshold, then T-consumers' consumption of the non zero-rated CP will be higher than that of S-consumers. The relationship between q_1^T and q_1^S when θ is below $d^S/(d^S + d^T)$ as well as the relationship between q_2^T and q_2^S when θ is above $d^S/(d^S + d^T)$ depend on the degree of content substitution, γ, and the difference $\tau^S - \tau^T$.

Second period: Content Providers' equilibrium strategies

Each CP's strategy is to decide whether to accept the ISP offer if any, in the fist stage of the second period, and then to set simultaneously and independently of the other CP, the price that maximizes its profit functions (see equation 2), given the ISP equilibrium strategy, and consumers' equilibrium quantities in the continuation game.

Starting with the second stage of the second period, it can be easily shown that $p_1 - p_2 = \frac{1+\gamma}{2+\gamma}(\alpha_1 - \alpha_2 + \bar{\tau})$. Therefore, not only the CP_1's equilibrium price will be higher than that of CP_2, but also the price difference will be greater than that under the neutral network regime. Repeating the exercise when the ISP zero-rates CP_2, and easy algebra shows that $p_1 - p_2 = \frac{1+\gamma}{2+\gamma}(\alpha_1 - \alpha_2 - \bar{\tau})$. Now, the price difference will be smaller than under the neutral network regime (and could even be reversed if the CPs had similar quality).

The economic intuition is as follows. When the ISP zero-rates CP_1 the price difference is driven by two forces in the same direction: First, CP_1 is the high-quality content; and second, there is an extra revaluation effect of its zero-rated content by eliminating the accounting from its consumption. All these forces add up to a CP_1's price higher than CP_2's price. However, when the ISP zero-rates CP_2, the price difference is driven by two opposite forces: First, again the CP_1 content is that of higher quality; and second, zero-rating CP_2 creates a competition effect caused by the fact that CP_1 will lower the price of its no zero-rated content to compete with the revalued content. Only if the content quality difference is high enough, then the CP_1's price will still be higher than CP_2's price.

When analyzing the relationship between each consumers' type equilibrium consumption from both CPs under a zero-rating offer to CP_1, we find that $q_1^m - q_2^m = \frac{1}{(1-\gamma)(2+\gamma)}[\alpha_1 - \alpha_2 + (2+\gamma)\tau^m - (1+\gamma)\bar{\tau}]$. Given that by $\tau^T \leq \bar{\tau} \leq \tau^S$, we can guarantee that $q_1^S > q_2^S$ and that this difference increases with the CPs' quality differentiation. However, the relationship between q_1^T and q_2^T is not clear. On the other hand, if ISP zero-rates CP_2, then $q_1^m - q_2^m = \frac{1}{(1-\gamma)(2+\gamma)}[\alpha_1 - \alpha_2 + (1+\gamma)\bar{\tau} - (2+\gamma)\tau^m]$ and we could only guarantee that $q_1^T > q_2^T$.

First period: Internet Service Provider's equilibrium strategies.

In the second stage of the first period, the ISP chooses its equilibrium network capacity and two-part tariff by solving a similar maximization problem to the one seen for the case of the neutral network, but with the objective function $\pi^{ISP} = \theta[H^T + \tau^T q_2^T] + (1-\theta)[H^S + \tau^S q_2^S] - c(\mu + \frac{1}{2}\mu^2)$

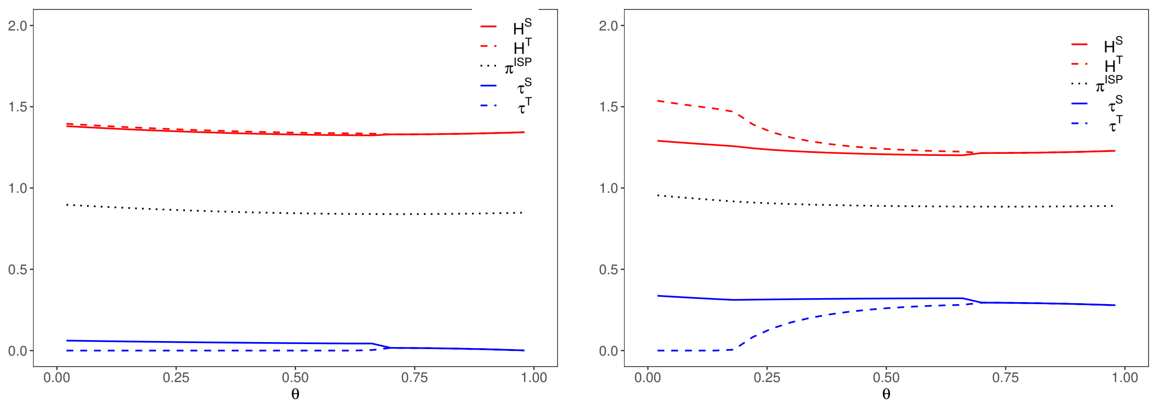
The choice of the network capacity is as the one under no zero-rating, that is, if only the S-consumers' incentive compatibility constraint and the T-consumers' participation constraint are binding, then $\mu = d^T/c - 1$. For the remaining variables of the ISP's strategy, we cannot derive again a closed-form solution. Therefore, we follow the same methodology than in the no zero-rating case. [20] Regardless of who is offered the zero-

[20] Following the same computational methodology than in the neutral network to solve the ISP maxi-

rating scheme by the ISP, we find similar results to those of the no zero-rating case: i) The ISP designs the interior equilibrium tariffs such that T-consumers' surplus is fully extracted, and S-consumers are indifferent between choosing any of the two-part tariffs; ii) $H^S \leq H^T$ and $\tau^S \geq \tau^T$; and iii) the ISP does not need to always set a menu of tariffs to extract the maximum profit.

Figure 2 shows the ISP's equilibrium hookup fees, per unit overage charges and profits as a function of the consumers' type distribution. The left panel displays the results when ISP zero-rates CP_1 and the right panel when ISP zero-rates CP_2.[21] Similarly to the neutral network regime, when the share of T-consumers is below a threshold $d^S/(d^S + dT)$, the ISP's equilibrium strategy is to offer a different two-part tariff to each type of consumer. Therefore, and again as in the no zero-rating case, as the differentiation between consumer types increases, the above threshold will also increase, and a single two-part tariff will require a higher proportion of T-consumers.

Figure 2: Equilibrium two-part tariff menu when the ISP zero-rates a CP



(a) ISP zero-rates CP_1 (b) ISP zero-rates CP_2

Notes: Two-part tariff menu has been computed for parameters $\alpha_1 = 1.5$, $\alpha_2 = 1$, $\gamma = 0.6$, $d^T = 0.1$, $d^S = 0.2$ and $c = 0.01$. Scale has been set equal for all plots.

Under the restricted network regime, we also have that T-consumers always consume from the two CP's content, and their demands decrease as they become more frequent in the market. On the contrary, S-consumers' demands increase as they become more scarce

mization problem, we find that the restrictions on the parameter values when the ISP offers zero-rating to CP_1 give us 1,210,052 instances for which interior and valid solutions exist. If instead to CP_1, the ISP offers zero-rating to CP_2, we find 1,221,243 combinations of market parameters for which there exits interior and valid solutions.

[21]Note that the patterns of the figures are similar in spite of γ=0.6 now. We have chosen this new value of γ for later comparisons.

in the market. However, the relationship between demands becomes more complex, as we will see in Section 6 (Equilibrium consumption patterns and network capacities).

5. Equilibria of the full game

Having derived the equilibrium strategies for the neutral network and the restricted network regimes, we analyze next the equilibria of the full game, where in the first stage of the first period the ISP has to decide whether to zero-rate a CP, and, in the first stage of the second period, the CPs have to decide whether to accept the ISP's offer, if any.

We start by studying whether the CPs will accept zero-rating offers. It is easy to prove that not only the price of a zero-rated content is higher than its price when it is not zero-rated, but also the consumption of its content is higher.

From equations (6) and (C.5) in Appendix C.1, and recalling that $\bar{\tau} = \theta\tau^T + (1-\theta)\tau^S$, some algebra shows that,

$$p_1^{ZR1} - p_1^{NZR} = \frac{1}{4-\gamma^2}[(1-\gamma)(2+\gamma)\bar{\tau}^{NZR} + \gamma\bar{\tau}^{ZR1}] \tag{14}$$

where the super-indexes NZR and $ZR1$ indicate when the equilibrium solution comes from no zero-rating or zero-rating policies, respectively. The right-hand side in (14) is always positive, therefore $p_1^{ZR1} > p_1^{NZR}$. In the same way we can show that $p_2^{ZR2} > p_2^{NZR}$.

In addition, from equations (4), (5), (C.1) and (C.2) in Appendix C.1, we can algebraically obtain that,

$$Q_1^{ZR1} - Q_1^{NZR} = \frac{1}{(1-\gamma^2)(4-\gamma^2)}[(1-\gamma)(2+\gamma)\bar{\tau}^{NZR} + \gamma\bar{\tau}^{ZR1}] \tag{15}$$

where $Q_1 = \theta q_1^T + (1-\theta)q_1^S$ is the (weighted) aggregate of consumers' demand of CP_1's content. The right-hand side in the above expression is always positive, hence $Q_1^{ZR1} > Q_1^{NZR}$. Equivalently, it is easily shown that under a zero-rating offer to CP_2, then $Q_2^{ZR2} > Q_2^{NZR}$. The following Lemma summarizes these results.

Lemma 2. *The equilibrium price of a zero-rated content is higher than its price when it is not zero-rated. Moreover, the equilibrium total demand of a zero-rated content is higher* *than its total demand when it is not zero-rated.*

Therefore, any CP will accept an (individual) offer of zero-rating by the ISP.[22] The introduction of a zero-rating offer, by eliminating the per unit overage charge, generates a double effect: first, the zero-rated CP has room to increase its price and subtract part of the consumers' surplus that the ISP no longer earns; second, the consumers' total payment by a unit of consumption from a zero-rated CP is lower than that in the no zero-rated case, so that consumers' consumption is higher in the first case than in the second. Then, any CP's profits are greater when it is zero-rated than when it is not. As a result,

Proposition 2. *It is a dominant strategy for CPs to accept a zero-rating offer.*

Focusing now on the ISP, we see that it has three strategies, which give rise to different continuation subgames: It can zero-rate either CP_1, or CP_2, or none. Given that the no zero-rating and zero-rating profit formulae are not handy at all, their analytical comparison becomes very cumbersome. A nice way to illustrate the ISP's profits under the above subgames is to show them graphically, for different combinations of the parameters, where we found that the degree of content substitutability, γ, is the most relevant parameter. The equilibrium calculation of the different subgames shows that the ISP's profits under zero-rating offers to CP_1 are always strictly Pareto-dominated by those under zero-rating offers to CP_2, and also by those under no zero-rating at all. Thus,

Result 2. *The ISP's strategy of zero-rating CP_1 is strictly Pareto-dominated by zero-rating CP_2 and also by no zero-rating at all. Therefore, at equilibrium the ISP will never offer zero-rating to CP_1.*

This result is similar to that found by Jeitschko et al. (2021), Lorenzon (2022), Gautier and Somogyi (2020) among others in models with homogeneous consumers: Without

[22]Easy algebra also shows that

$$Q_j^{ZRi} - Q_j^{NZR} = \frac{1}{(1-\gamma^2)(4-\gamma^2)}[(1-\gamma)(2+\gamma)\bar{\tau}^{NZR} - (2-\gamma^2)\bar{\tau}^{ZRi}].$$

Hence, the consumption by the two consumer types of the no zero-rated content could be either higher or lower than its consumption when it is no zero-rated. Hence, it might be possible that the no zero-rated CP is better off under zero-rating than under no zero-rating at all of any CP.

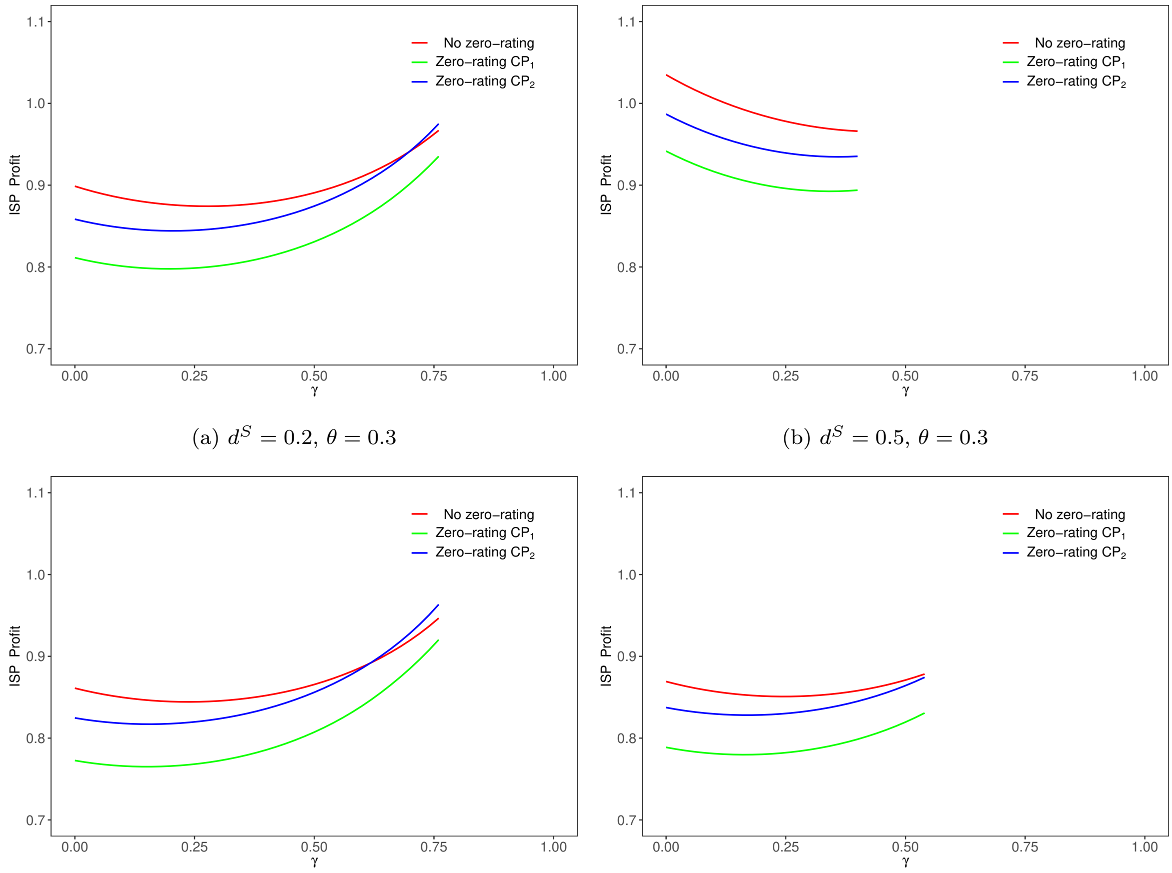
monetary transfer for zero-rating, the ISP will offer zero-rating to the low-quality content provider, if the degree of content substitution is sufficiently high. Clearly, if the ISP zero-rates CP_2's content, the ISP's total profit will be the sum of the hookup fees and the overage charges from consumers of $CP1$. When content presents a low substitution, demands are quite inelastic and even the lower quality CP_2 has relatively large demands. Then, the ISP's profit loss from eliminating the overage charge of CP_2's demand is large enough to prevent the ISP from offering zero-rating. On the contrary, when γ is high enough, lowering the cost to consumers of CP_2 may bolster demand and allow the ISP to more than recoup lost profits through hookup fees.

However, as it will become clear below consumers' heterogeneity adds new angles to the analysis. The ISP provides a unique network capacity and the traffic generated on the network is constrained by the consumers' sensitivity levels of network congestion. Therefore, the decision to patronize one CP over the other will also depend on the relative mass of the two types of consumers and their network speed preferences.

Figure 3 compares the ISP's profits of no zero-rating (red line), zero-rating CP_1 (green line), or zero-rating CP_2 (blue line) as a function of the degree of content substitutability, γ, for different values of d^S (consumers' heterogeneity) and θ (the proportion of T-consumers)[23]. In all cases the green line is below the others (i.e., zero-rating CP_1 is dominated by the other cases). As Figures 3.a and 3.c reveal, when there is almost no consumer differentiation ($d^T = 0.1$ and $d^S = 0.2$), the ISP will prefer zero-rating CP_2 if γ is above a certain threshold value, otherwise the ISP will prefer no zero-rating. This threshold decreases as the proportion of T-consumers increase. On the other hand, when there is a high consumer differentiation ($d^T = 0.1$ and $d^S = 0.5$), the ISP will prefer no zero-rating independently of γ (see Figures 3.b and 3.d). In summary, Figure 3 reveals that the market conditions that govern the ISP's choice (no zero-rating vs. zero-rating CP_2) to extract the maximum surplus exhibit a complex behaviour with respect to the parameters of the model and deserve a more detailed analysis.

[23]Recall that θ is the proportion of content consumers with less sensitivity to the network waiting time or latency. According to OECD (OECD, 2024), approximately 30% of Internet use is mainly made by business activity, meanwhile approximately 70% is for entertainment. Moreover, the former tolerates more latency than the latter, 100ms and 50ms, respectively. Regarding our model, the above data translates to a d^S double than d^T, and θ around 30%.

Figure 3: ISP's profits under zero-rating and no zero-rating strategies



(a) $d^S = 0.2$, $\theta = 0.3$

(b) $d^S = 0.5$, $\theta = 0.3$

(c) $d^S = 0.2$, $\theta = 0.7$

(d) $d^S = 0.5$, $\theta = 0.7$

Notes: ISP's profits are computed for parameters $\alpha_1 = 1.5$, $\alpha_2 = 1$, $d^T = 0.1$ and $c = 0.01$. Scale has been set equal for all plots.

To better explain the ISP's equilibrium choice, Figure 4 displays four panels, where each one of them shows the ISP's choice as a function of parameters γ and θ. Namely, the panels display the equilibrium joint strategies followed by the ISP: no zero-rating vs. zero-rating CP_2, and either a menu of two-part tariff or a single two-part tariff. In other words, whether the ISP makes use at equilibrium of either the two discriminatory strategies, one of them, or none, as a function of the market conditions.

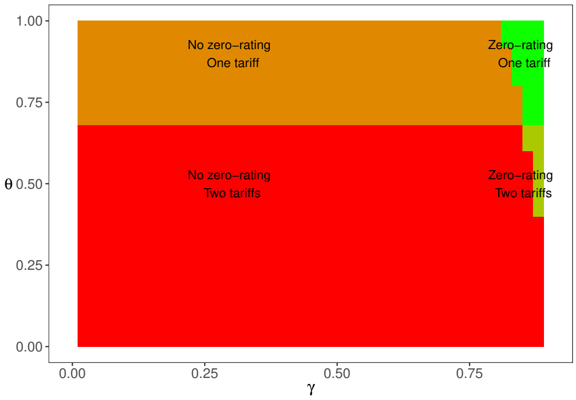
The red and brown areas correspond to those pairs (γ, θ) such that the ISP chooses at equilibrium no zero-rating, and the light green and dark green areas to those pairs such that ISP chooses zero-rating CP_2. Moreover, brown and light green areas correspond to those markets where the ISP uses a single two-part tariff, meanwhile red and dark green areas to those markets where a two-part tariff menu is used. The top-hand row panels correspond to the case where there is a low quality differentiation of CPs ($\alpha_1 = 1.2$, quite close to $\alpha_2 = 1$) and the bottom-hand row panels correspond to a high-quality differentiation ($\alpha_1 = 1.5$). On the other hand, in the left-hand panels there is a low consumers' differentiation ($d^T = 0.1$ and $d^S = 0.2$), meanwhile in the right-hand panels there is a high consumers' differentiation ($d^T = 0.1$ and $d^S = 0.5$). The analysis of the four panels reveals that the ISP's equilibrium decision of zero-rating CP_2 depends on the degree of substitutability between contents and the consumer types' distribution, which, in turn, are related with the differentiation of both CPs and consumers.[24] However, the ISP's equilibrium choice of either a menu of two-part tariff or a single two-part tariff depends only on consumer factors: Their type distribution and their speed preference. In particular on the relationship between $(1-\theta)d^S$ and θd^T, that is, on θ higher or smaller than $d^S/(d^S + d^T)$.

At equilibrium, the ISP will choose zero-rating CP_2 when the degree of substitutability, γ, and the proportion of the T-consumer type, θ, are beyond certain thresholds, as the green areas in the top-right corner in the panels reveals. The thresholds are not independent of each other, the higher the threshold on γ, the lower the threshold on θ. As the consumers' differentiation increases, the pairs (γ, θ) for which the ISP offers zero-rating to CP_2 are reduced. The green areas in the left-hand panels of Figure 4 are bigger than the green areas in the right-hand panels. Finally (as already revealed in Figures 1 and 2), the ISP will set a single two-part tariff when the proportion of T-consumer is above $d^S/(d^S + d^T)$; thus, the higher the difference between consumers' speed preferences the higher the proportion of T-consumer needed to implement a single two-part tariff.

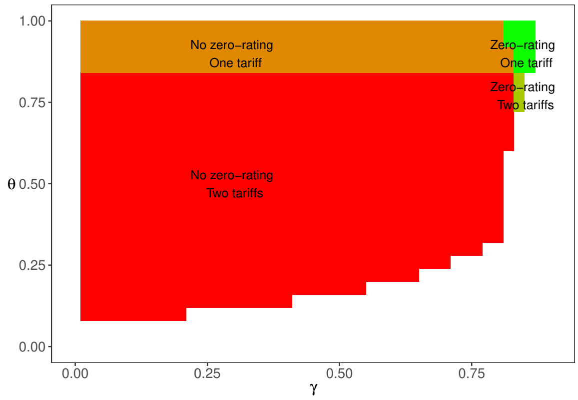
From a mirror angle, relating the use of the ISP's discrimination tools with the joint heterogeneity of both CPs and consumers, we find that the highest ISP's joint discrimination to both CPs and consumers will take place when the CP's differentiation is high and the consumers differentiation is not too high (see Figure 4.c). In this case, the ISP will set a two-part tariff menu that seeks to change the demand patterns to internalize the congestion disutility of S-consumers, and will zero-rate CP_2 as a content revaluation tool, changing even more consumption patterns. On the contrary, the smallest joint discrimination by the ISP will occur when CP's differentiation is small and the consumers

[24]As stated in sections 3 and 4, for some combinations of the market parameters there is neither a interior solution nor a valid solution. In this cases the ISP's maximization problem has not a solution and the area in the panel remains white.

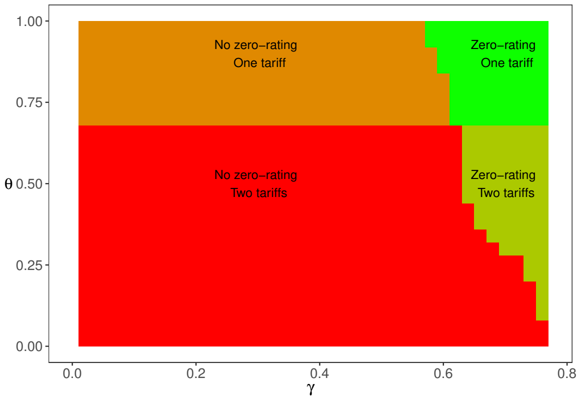
Figure 4: The ISP's equilibrium discrimination strategies: single tariff versus tariff menu, and no zero-rating vs. zero-rating CP_2



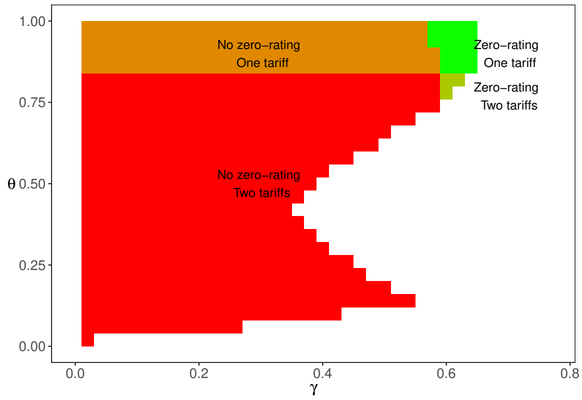
(a) $\alpha_1 = 1.2$; $d^S = 0.2$



(b) $\alpha_1 = 1.2$; $d^S = 0.5$



(c) $\alpha_1 = 1.5$; $d^S = 0.2$



(d) $\alpha_1 = 1.5$; $d^S = 0.5$

Notes: The ISP's profits are computed for parameters $\alpha_2 = 1$, $d^T = 0.1$ and $c = 0.01$. Scale has been set equal for all plots.

differentiation is high (see Figure 4.b). These results, however, have to be qualified by the degree of content substitutability. Specifically, take, for example, $\theta = 0.5$ and $d^S = 0.2$ (left-hand panels in Figure 4). At these values of θ and d^S, the ISP's equilibrium strategy is to choose a two-part tariff menu, but if γ is below a certain threshold, no zero-rating will take place. As the CP's differentiation increases the threshold on γ needed for zero-rating is smaller and keeps decreasing as CP's differentiation does so. However, note that the more heterogeneous are consumers the lesser the equilibrium zero-rating offers to CP_2. Therefore, necessary conditions for equilibrium with zero-rating offers are high substitution between contents, heterogeneous CPs, and consumers not too heterogeneous. We summarize the subgame perfect equilibria of the full game in the following Result.

Result 3. *The subgame perfect equilibria of the game under interior solutions, and their characterization are,*

1. *There are four types of subgame perfect equilibria. Each of them is given by the ISP's play of the two discriminatory strategies (zero-rating and tariff menu), of one of them (either zero-rating or tariff menu), or of none (single two-part tariff and no zero-rating), and its existence depends of the model parameters.*
2. *There exists a threshold frontier $(\hat{\gamma}, \hat{\theta})$ such that for pairs $(\gamma,\ \theta)$ on and above this frontier, the ISP's equilibrium strategy is always to zero-rate CP_2. For pairs $(\gamma,\ \theta)$ below the threshold frontier, the ISP will play no zero-rating.*
3. *The threshold frontier $(\hat{\gamma}, \hat{\theta})$ depends on the heterogeneity of CPs as well as on consumers' differentiation. Thus, a high level of substitution between contents, heterogeneous CPs, and consumers not too heterogeneous are necessary conditions for an equilibrium zero-rating offer to CP_2.*
4. *In equilibrium, the ISP will choose a single two-part tariff, if θ is beyond the threshold $\bar{\theta}$, that makes $(1-\bar{\theta})d^S = \bar{\theta}d^T$. This threshold is independent of γ, but it depends on the difference between consumers' speed preferences. For θ below $\bar{\theta}$, the ISP will set a tariff menu. The greater the heterogeneity of consumers, the higher the threshold $\bar{\theta}$.*

The economic intuition of the frontier is clear: What makes profitable zero-rating by the ISP is the change in consumption patterns, which allows the ISP to internalize the S-consumers' congestion externality. Therefore, the change of CPs' relative prices must provide the necessary incentive and the S-consumers' externality be not too big. A high substitution level of CPs' content is necessary to change consumers' consumption induced by the change of relative prices due to zero-rating strategies. Also, to guarantee the CPs' separation, heterogeneity of their content is needed, because, otherwise, CPs will set very similar prices and consumers will purchase very similar content quantities. The two conditions are necessary, because, if content is substitutable but homogeneous, it will be very difficult to redirect consumption patterns by zero-rating strategies. On the contrary, if content is heterogeneous but not substitutable, there is no room to separate consumption. Finally, consumer differentiation must be not too high because otherwise the S-consumer congestion externality would be too high to profitably redirect the traffic flow. [25]

[25] To fulfil the incentive and participation constraints of the ISP maximization problem.

In addition, there is an inverse relationship between the $\hat{\gamma}$ and the $\hat{\theta}$ of the threshold frontier. Namely, the content substitutability threshold for γ decreases as the proportion of T-consumers –θ– increases. This means that if the share of T-consumers is high enough, then γ does not need to be very large for the ISP to choose zero-rating CP_2. This is because, by equations (4), (C.7) and (C.9), the T-consumers always consumes from both content, while under no zero-rating, by equation (5), the S-consumer may sometimes prefer to consume only the content she values most, CP_1. Thus, the ISP has an incentive to zero-rate CP_2. By doing that, the ISP increases the consumption of content and therefore the possibility of extracting more surplus.

Finally, consumers' differentiation on speed preferences fixes the θ from which the regime of two-part tariff menu collapses to a single two-part tariff. This θ is the one from which on the S-consumers' congestion externality is completely internalized by the ISP, under both regimes: neutral network and restricted network. As S-consumers' consumption increase as θ increases, there is a θ from which on it will surpass T-consumers' consumption and there in no further need to discriminate between the two types of consumers. Thus, when the share of T-consumers is high enough (depending of the consumers' differentiation), the ISP will zero-rate CP_2 and set a single two-part tariff.

A deeper analysis of the two types of consumers' consumption is offered in the next Section.

6. Equilibrium consumption patterns and network capacity

The ISP wants to internalize the network congestion externality of S-consumers. This can be achieved in first place through a tariff menu, changing the content relative prices and thus managing the demand patterns, without increasing the network capacity. However, this discrimination device does not achieve a demand increase with respect to the single tariff[26] and, therefore, a further increase in ISP's profits. What does zero-rating contribute to ISP's profits? The answer is clear: The possibility of changing content relative prices even more and then increasing demands by redesigning consumption pat-

[26] As already mentioned, the comparison between aggregate total consumption under a single two-part tariff and a menu of tariffs is offered in detail in Appendix B.2

terns. As already mentioned, to do that it is necessary that content is highly sustitutable and heterogeneous. To understand the mechanisms at work, Figure 5 illustrates the equilibrium consumption patterns of both consumer types as a function of the proportion of T-consumers in the market. The panel in the left hand side shows the no zero-rating case (5.a) and that in the right hand side (5.b) when the ISP zero-rates CP_2.

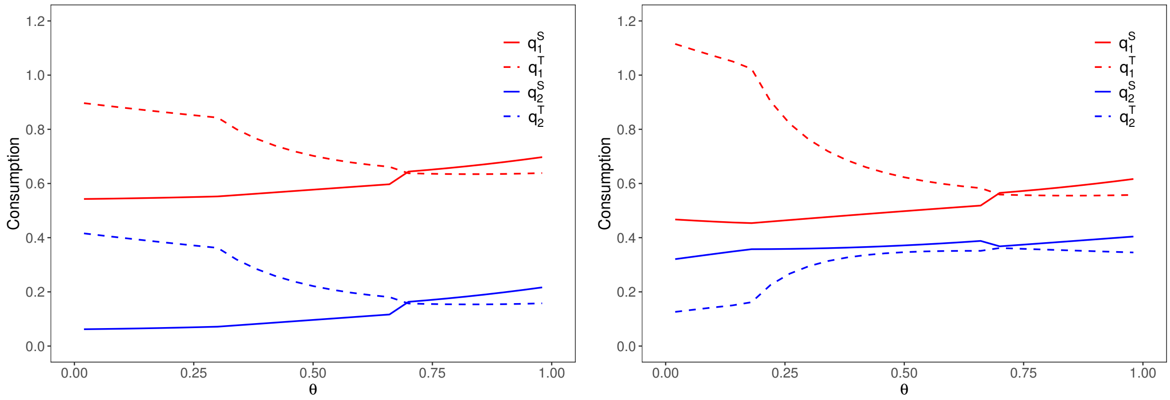
In the no zero-rating case, we observe that both types of consumers consume more from CP_1 than from CP_2. This can be easily seen in the left hand side panel, where the solid (dashed) red line is above the solid (dashed) blue line. We also observe that when the proportion of T-consumers is high enough, they consume less from any CP than S-consumers: In some point, the solid lines cross upwards the dashed lines. Interestingly enough, and as in the classical models on quality-price of Mussa and Rosen (1978), a partial exclusion effect shows up: When consumers are quite differentiated and the proportion of S-consumers is quite high, they do not consume CP_2's content at all as they prefer to focus on the most valued content.[27] In fact, figures B.12.a and B.12.b in Appendix B.2 show that the equilibrium aggregate consumption under a two-part tariff menu is always smaller than or equal to that of a single two-part tariff. Consequently, the average level of congestion, $\mu - \theta(q_1^T + q_2^T) - (1-\theta)(q_1^S + q_2^S)$, is at least as higher under a two-part tariff menu than under a single two-part tariff.

Under zero-rating CP_2, the relationship between consumers' demands becomes more complex. By equations (12) and (13), and as stated in section 4, the T-consumers' consumption from CP_1 is always higher than that from CP_2: In the right hand panel in Figure 5 the dashed red line is above the dashed blue line, and the same happens with S-consumers' demands. However, the relationship between the two types of consumers' demands depends now, among other parameters, on the proportion of T-consumers and the degree of content substitution γ. Only for a proportion of T-consumers above a threshold, S-consumers consume more from CP_1 than T-consumers, otherwise the reverse is true; and S-consumers' demand of CP_2 is always higher than that of T-consumers. As expected, and as Figure 5.b illustrates, under a zero-rating offer to CP_2 the two-types of consumers always consume from both CP's contents, independently of marker parameters, and the consumption of CP_2's content is increased. T-consumers' total demand decrease

[27]In fact, under no zero-rating and $\gamma = 0.2$, S-consumers do not consume from CP_2 when they are the majority, i.e., for small values of θ.

as they become more frequent in the market (sum of dashed lines), and S-consumers' total demand increases as they become scarcer in the market (sum of solid lines).

Figure 5: T-consumers and S-consumers' demands of CP_1 and CP_2's contents under no zero-rating and zero-rating CP_2.



(a) No zero-rating

(b) Zero-rating CP_2

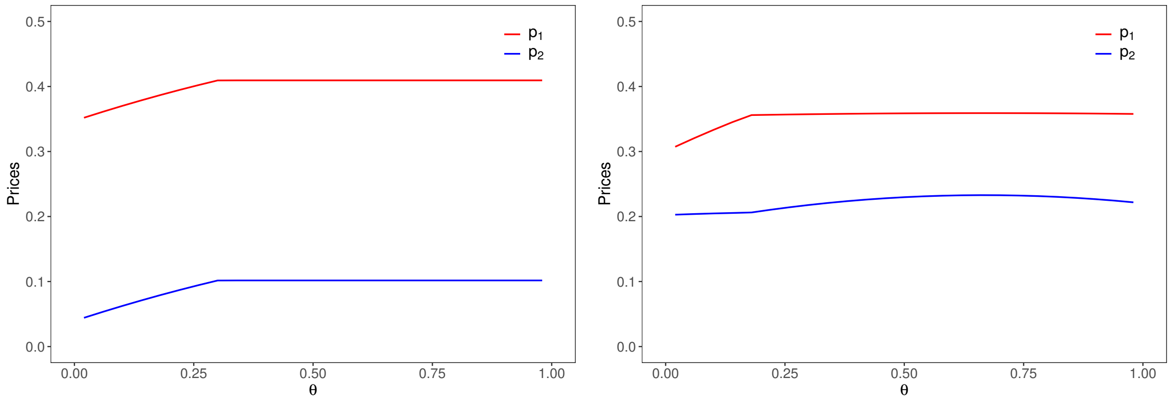
Notes: Consumers' equilibrium consumption has been computed for parameters $\alpha_1 = 1.5$, $\alpha_2 = 1$, $\gamma = 0.6$, $d^T = 0.1$, $d^S = 0.2$ and $c = 0.01$. Scale has been set equal for all plots.

Result 4. *At the equilibrium where the ISP chooses no zero-rating, T-consumers always consume content from both CPs, but S-consumers may consume content from either CP_1 or both CPs. At the equilibrium where the ISP zero-rates CP_2, both types of consumers consume from both CP's content.*

Given the parameter values of Figure 5, we can relate its panel b with Figure 4.c. In fact, if we set the value $\gamma = 0.6$ in the later, the equilibrium strategies of the ISP will be zero-rating CP_2 when θ is higher than or equal to 0.67 and will set a single two-part tariff. For smaller values of θ, the ISP will only practice price-discrimination through a tariff menu, showing the above mentioned relationship between both parameters in the equilibrium threshold frontier. Thus, only when content substitution is high enough zero-rating will be implemented. Zero-rating can be offered by the ISP either under a single two-part tariff or jointly with a tariff menu. For example, if $\gamma = 0.7$ and θ is lower than 0.67, the ISP will use both a tariff menu and zero-rating, as figure 4.c illustrates.

Note that under zero-rating (see Figure 5.b) q_2^S is higher than q_2^T. The change in the consumption pattern follows the incentives of the PCs' changes on relative prices induced by zero-rating CP_2. That is, the total price of CP_2's content for consumers is lower due to

Figure 6: CP_1 and CP_2 prices under no zero-rating and zero-rating CP_2.



(a) No zero-rating (b) Zero-rating CP_2

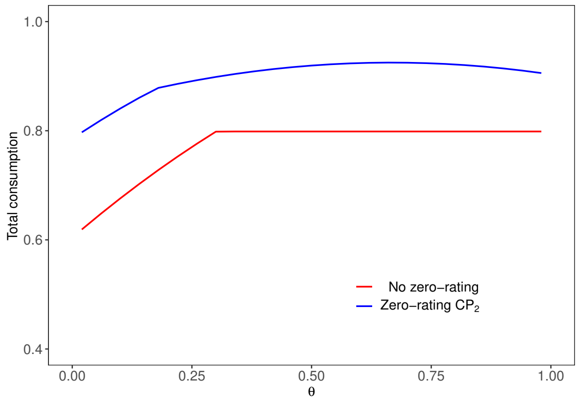
Notes: CP's equilibria prices have been computed for parameters $\alpha_1 = 1.5$, $\alpha_2 = 1$, $\gamma = 0.6$, $d^T = 0.1$, $d^S = 0.2$ and $c = 0.01$. Scale has been set equal for both plots.

the removal of tau, and although CP_2' price goes up the resulting price is still lower than without zero-rating. Consequently, CP_1' price goes down to be competitive (see Figure 6). As a result, aggregate consumption increases with respect no zero-rating. By zero-rating the less valued CP, the ISP guarantees that both consumers will buy from the two CPs, increasing consumption and hence decreasing the network flow and increasing network delays. However, the total price reduction of CP_2's content compensates the network delay costs and, therefore, in the new consumption pattern, the more speed demanding consumers buy more from CP_2 and the less speed demanding consumers may buy more from CP_1, and less form CP_2 (being like an inferior good for this consumers' type). Figure 7 clearly depicts the increase in the equilibrium aggregate consumption under zero-rating compared to no zero-rating. Clearly, the average network congestion increases, as compared to a two-part tariff menu, but the two-types of consumers buy more content because the new content prices compensate the disutility of the network delay cost.

In line with Choi and Kim (2010), and as in Lorenzon (2022) our results show that a zero-rating contract to the least attractive CP reduces the ISP's incentives to invest in network capacity. More specifically, zero-rating plans allow the ISP to target consumers' consumption profiles and offer a more effective method of price discrimination.

It is important to remark that since the optimal μ only depends on the ISP's cost associated to the network capacity (c) and the marginal revenue of the lowest consumers'

Figure 7: Equilibrium aggregate consumption under no zero-rating and zero-rating CP_2



Notes: (Weighted) aggregate consumption has been computed for parameters $\alpha_1 = 1.5$, $\alpha_2 = 1$, $\gamma = 0.6$, $d^T = 0.1$, $d^S = 0.2$ and $c = 0.01$, and scale has been set equal for both plots.

type preference for network speed (d^T), then the ISP's optimal network choice will be the same under the single two-part tariff, the tariff menu, the zero-rating scheme, and the neutral regime. Therefore, the analysis is useful both in the short-term and long-term. As long as there are consumers with different network speed preferences, and CPS of different quality, the ISP may use the redirection of traffic flows on the network to increase its profits. Therefore, both the tariff menu and the zero-rating offers are both short-term and long-term discriminatory strategies, which allow the ISP to internalize negative congestion externalities without increasing its network capacity.

7. Social Welfare

Consider now Social Welfare under the different equilibria. We wish to assess how the zero-rating equilibrium affects welfare compared to the equilibrium without zero-rating, when consumers have heterogeneous preferences for the network speed and CPs differ in quality. Define Social Welfare as the (weighted) sum of the surplus of the two types of consumers, CPs' profits and ISP's profit:

$$SW = \theta U^T + (1-\theta)U^S + \pi_1^{CP} + \pi_2^{CP} + \pi^{ISP}.$$

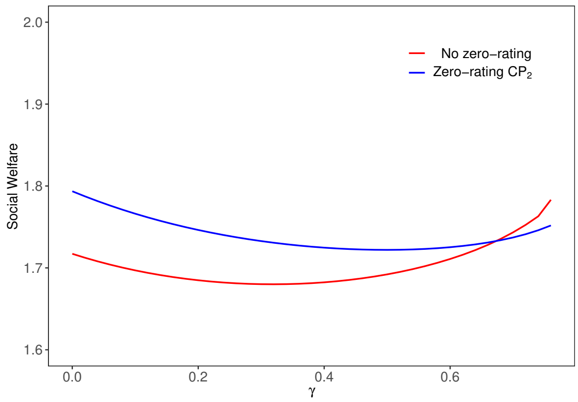
Simplifying the above expression, we find that

$$
\begin{aligned}
SW &= \theta u(q_1^T, q_2^T) + (1-\theta)u(q_1^S, q_2^S) + \\
&\quad (\theta d^T + (1-\theta)d^S)[\mu - \theta(q_1^T + q_2^T) - (1-\theta)(q_1^S + q_2^S)] - \\
&\quad c(\mu + \frac{1}{2}\mu^2).
\end{aligned}
$$

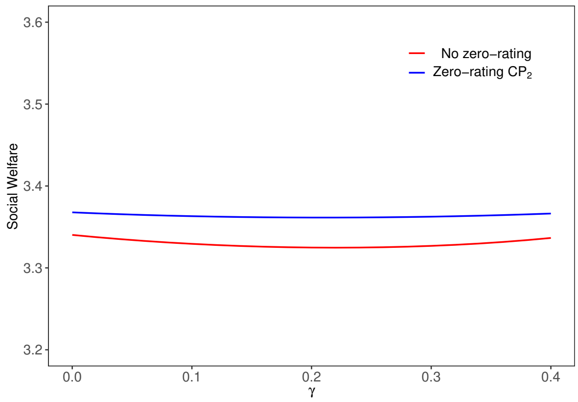
Given the complex calculation and to highlight our results, we show them graphically. Thus, Figure 8 compares the Social Welfare achieved under no zero-rating and when the ISP zero-rating CP_2. Each panel in Figure 8 corresponds to different consumers' type distributions and degrees of consumers' heterogeneity. Thus, the top-hand panels display Social Welfare when the majority of consumers are of S-type ($\theta = 0.3$), the left-top-hand panel (Figure 8.a) when consumers are almost homogeneous ($d^T = 0.1$ and $d^S = 0.2$), and the right-top-hand panel (Figure 8.b) when consumers are more heterogeneous ($d^T = 0.1$ and $d^S = 0.5$). Figures 8.c and 8.d repeat the same scheme when the majority of the consumers are of T-type ($\theta = 0.7$).

Figure 8 reveals that the ISP's zero-rating offer to CP_2 generates a greater Social Welfare than no zero-rating at all, depending on consumers' differentiation and the degree of substitutability. When consumers are quite homogeneous ($d^S = 0.2$) zero-rating CP_2 will increases Social Welfare if contents are not highly substitutable. However, if consumers are quite differentiated ($d^S = 0.5$), Social Welfare is higher under zero-rating, independently of γ. To get some economic intuition on the above results, Figure 9 shows the decomposition (among consumers, CPs, and the ISP) of the Social Welfare corresponding to Figure 8.c, when consumers are quite homogeneous ($d^T = 0.1$ and $d^S = 0.2$); and Figure 10 shows the same decomposition corresponding to Figure 8.d, when consumers are quite heterogeneous ($d^T = 0.1$ and $d^S = 0.5$). In both cases the majority of the consumers are of T-type ($\theta = 0.7$). The four panels in Figures 9 and 10 correspond to the four different agents in the market and show their benefits under no zero-rating (blue lines) and zero-rating to CP_2 (red lines). More precisely, the top-left panel displays the S-consumer' utility –recall that at equilibrium the ISP extracts the entire T-consumer surplus; the top-right panel shows the ISP's profits; the bottom-left panel the CP_1 profits; and the

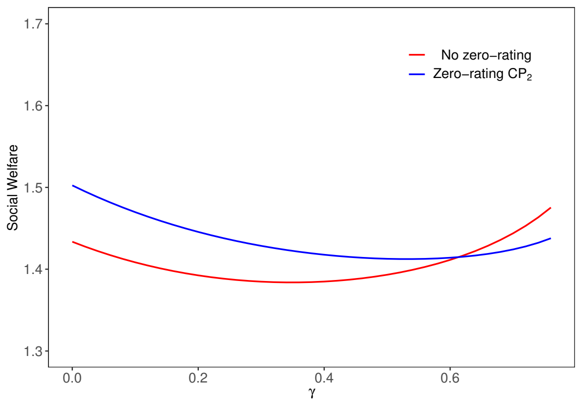
Figure 8: Social Welfare as a function of the consumers' type distribution and their preferences for network speed



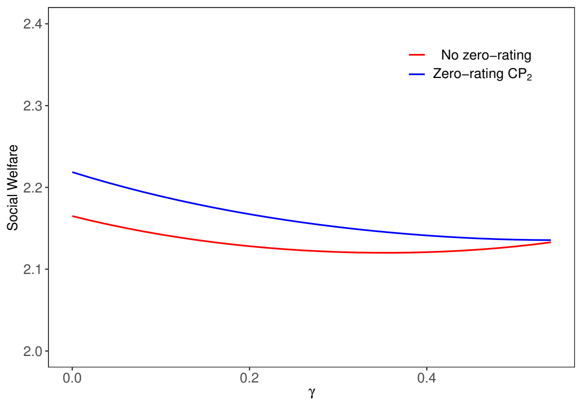
(a) $d^S = 0.2, \theta = 0.3$



(b) $d^S = 0.5, \theta = 0.3$



(c) $d^S = 0.2, \theta = 0.7$



(d) $d^S = 0.5, \theta = 0.7$

Notes: Social welfare has been computed for parameters $\alpha_1 = 1.5, \alpha_2 = 1, d^T = 0.1$ and $c = 0.01$. Scales are different in each Figure.

bottom-right panel the CP_2 profits.

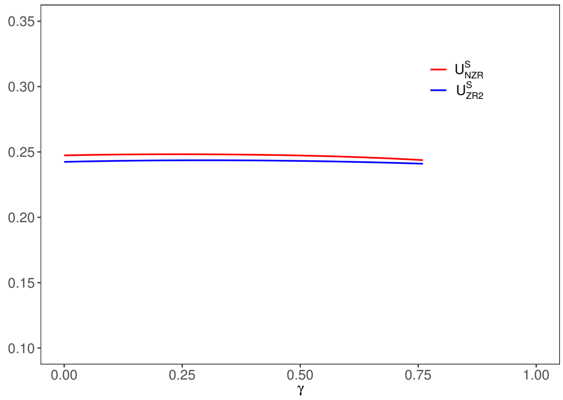
When consumers are quite homogeneous (Figure 9), S-consumer and CP_1 will be better off under no zero-rating. On the contrary, as we already knew from Figure 3, the ISP will be better off under zero-rating if the degree of substitution between CPs' content is high enough; however, as stated in Proposition 2, CP_2 is better off accepting zero-rating. In aggregation, for values of γ below certain threshold the extra profits that CP_2 obtains when she is zero-rated compensate the extra benefits that the other agents obtain under no zero-rating. Thus, the maximum Social Welfare occurs when the ISP zero-rates CP_2 as shown in figure 8.c. As γ increases, the extra profits that CP_2 obtain when she is zero-rated decrease and the extra profits that CP_1 obtain under no zero-rating increase. Consequently, in aggregation the above mentioned compensation is no longer in place and the maximum Social Welfare occurs under no zero-rating. Note that the agents more affected by the ISP strategy are the CPs, and, in fact, their profits are what ultimately determines when Social Welfare is maximized. However, as Figures 4.a and 4.c show, the ISP may or not find it profitable zero-rating CP_2, depending on the value of θ, the associated γ in the equilibrium frontier, and the CPs' differentiation.

When consumers are quite heterogeneous (Figure 10) we again observe that the CPs are the market agents more affected by the ISP strategy, but there are two main differences with respect to the more homogeneous consumers' case. First, the CP_1 extra profits under no zero-rating remains almost constant and independent of the degree of content substitution; second, even for a high degree of content substitution, the CP_2 extra profits when it is zero-rated remain high enough. As a result, Social Welfare is always maximized under zero-rating CP_2. But now, as Figures 4.b and 4.d show, zero-rating offers to CP_2 only take place for a very limited set of parameter values. We summarize these findings in the following Result.

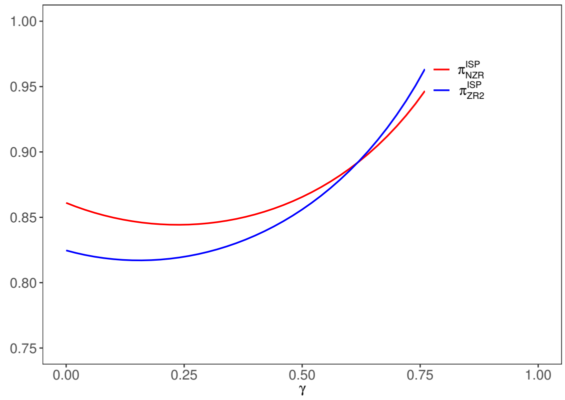
Result 5. *When consumers are quite homogeneous, zero-rating offers to CP_2 by the ISP will generate greater social welfare than no zero-rating at all, unless the CPs' degree of content substitution is very high. However, consumers and CP_1 are worse off under zero-rating CP_2.*

When consumers are quite heterogeneous instead, Social Welfare is higher under zero-

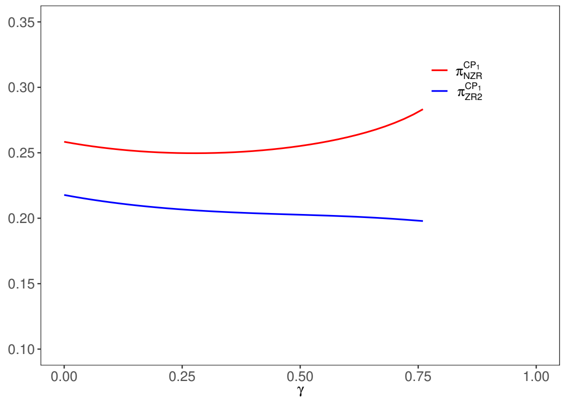
Figure 9: Social Welfare decomposition. Low consumers' differentiation



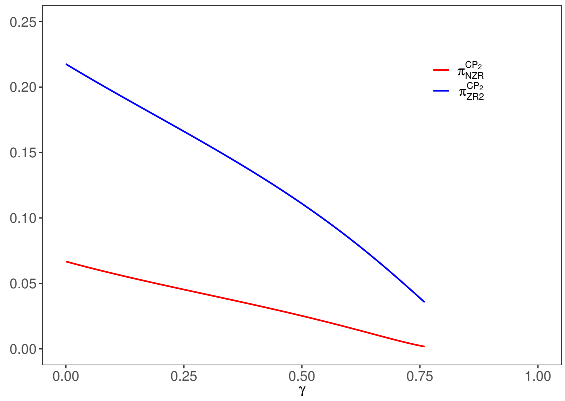
(a) Consumers utility



(b) ISP profits



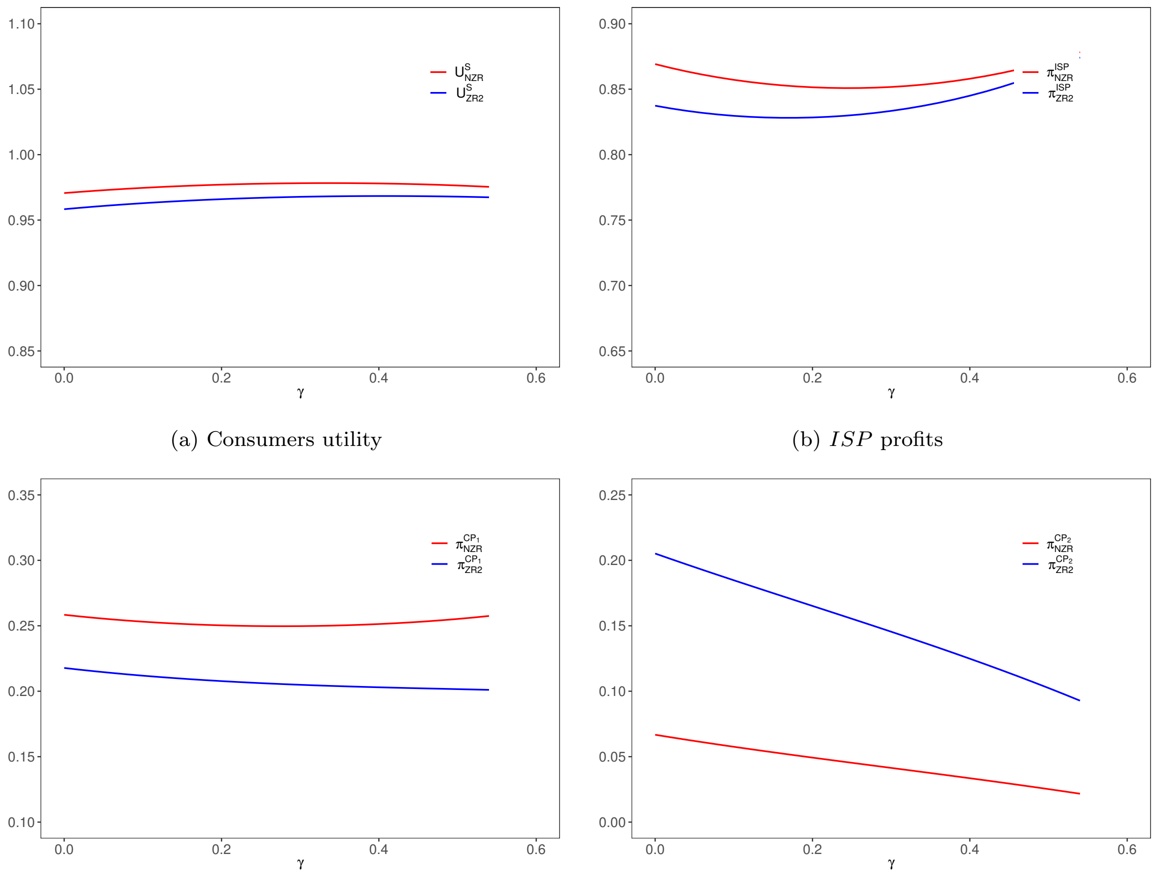
(c) CP_1 profits



(d) CP_2 profits

Notes: Profits scale are not equal for all figures, but the range of the scale is 0.25 in the four panels. Therefore, the differences between the two schemes are comparable across the four panels. Social welfare has been computed for parameters $\alpha_1 = 1.5$, $\alpha_2 = 1$, $d^T = 0.1$, $d^S = 0.2$, $c = 0.01$ and $\theta = 0.7$.

Figure 10: Social Welfare decomposition. High consumers' differentiation



(a) Consumers utility

(b) ISP profits

(c) CP_1 profits

(d) CP_2 profits

Notes: Profits scale are not equal for all figures, but the range of the scale is 0.25 in the four panels. Therefore, the differences between the two schemes are comparable across the four panels. Social welfare has been computed for parameters $\alpha_1 = 1.5$, $\alpha_2 = 1$, $d^T = 0.1$, $d^S = 0.5$, $c = 0.01$ and $\theta = 0.7$.

rating CP_2, independently of the degree of substitution between contents. Again, consumers and CP_1 are worse off under zero-rating CP_2.

Schnurr and Wiewiorra (2018) find that the combination of data caps with zero-rating can be used to discriminate between different types of consumers depending on the level of heterogeneity. They find that when consumer groups are heterogeneous in their valuation of zero-rated contents, they benefit from this practice. Otherwise, zero-rating is detrimental to them. Unlike Schnurr and Wiewiorra (2018) we find that the use of zero-rating services with either a menu of tariffs or under a single two-part tariff is a discrimination tool which is always detrimental to consumers.

8. Conclusions

In this paper we have analyzed the effects of zero-rating in a monopoly market where the ISP can offer a menu of tariffs to two types of consumers. The consumers, type Turtle (T) and type Speedy (S), are heterogeneous in their preference for network speed, or in other words, in how network congestion affects their demands of online content. Also, the CPs differ in the quality of their content. The monopolist could alleviate the congestion problem by increasing its network capacity. However, the ISP does not want to do this because it comes at a cost and looks for the strategies that allow it to maximize its profits without expanding its capacity. In this regard, the ISP has two powerful discrimination tools in the market to internalize the congestion externality: two-part tariff menus, which may be set by the ISP to consumers as a screening tool to change the consumption pattern, and zero-rating services to revalue content and discriminate between CPs, to redirect consumption even more.

We show that the subgame perfect equilibria with zero-rating are separated from those without zero-rating by a frontier, which depends on the degree of substitution between contents and the market shares of the consumers' types. In addition, we have found that the greater the substitution between content, the greater the heterogeneity between content providers, and the greater the homogeneity between consumers, the greater the range of existence of zero-rating equilibria. Moreover, and in particular, when consumers are quite heterogeneous zero-rating offers only occur for a very limited set of parameter values.

Finally, a key issue in zero-rating regulation is its impact on Social Welfare. As Inceoglu and Liu (2019), we found that zero-rating may improves Social Welfare under consumers heterogeneity. When consumers are more homogeneous, Social Welfare only increases with zero-rating when there is not too high substitution between CPs. However, and unlike Schnurr and Wiewiorra (2018), we find that the use of zero-rating services with either a menu of tariffs or under a single two-part tariff is a discrimination tool which is always detrimental to consumers. Thus, both types of consumers and the no zero-rated CP are always worse off under zero-rating. In this regard, we approach Preta and Peng (2020) to point out that, for the policymaker it would be interesting to see how the surplus is redistributed between ISPs and CPs by the use of these practices and their potential effects on competition and its consequences for antitrust laws.

Our findings have some implications for regulatory policy. Practices that infringe on net neutrality not only cause worries in upstream markets for content provision, but may also increase the ability of a dominant ISP to extract greater surplus from consumers. In June 2022, BEREC (2020) updated the guidelines on zero-rating services following the case law of the European Court of Justice (ECJ) in 2021. BEREC (2022) includes a strict ban on zero-rating practices. Different reasons for the ban on zero-rating have been offered ranging from distortion of competition between content providers, barriers to entry in content markets, the hindering innovation and development of services, etc. Our study adds to these reasons by showing that $ISPs$ do not provide the capacity needed to contribute to the efficient use of network resources and to an optimization of the overall quality of transmissions.

Hence, our results support BEREC's decision to ban zero-rating practices in the EU. With such a decision, the regulatory body prevents traffic distorting practices that harm the end user as well as anti-competitive effects on smaller content providers.

One of the limitations of this study is that it restricts the market to a single ISP. Future research should study the effect of competition between Internet service platforms on the use of these two price discrimination tools.

References

Armstrong, M. (1996). Multiproduct nonlinear pricing. *Econometrica*, 64(1):51–75.

BEREC (2016). Guidelines on the implementation by national regulators of european net neutrality rules. https://www.berec.europa.eu/en/open-internet/introduction.

BEREC (2020). Guidelines on the implementation of the open internet regulation. https://www.berec.europa.eu/en/open-internet/introduction.

BEREC (2022). Guidelines on the implementation of the open internet regulation. https://www.berec.europa.eu/en/open-internet/introduction.

Bourreau, M., Kourandi, F., and Valletti, T. (2015). Net neutrality with competing internet platforms. *The Journal of Industrial Economics*, 63(1):30–73.

Choi, J. P. and Kim, B.-C. (2010). Net neutrality and investment incentives. *The RAND Journal of Economics*, 41(3):446–471.

Choné, P. and Linnemer, L. (2020). Linear demand systems for differentiated goods: Overview and users's guide. *International Journal of Industrial Organization*, 73:1–24.

Economides, N. and Hermalin, B. E. (2015). The strategic use of download limits by a monopoly platform. *The RAND Journal of Economics*, 46(2):297–327.

Eisenach, J. A. (2015). The economics of zero rating. https://www.berec.europa.eu/en/open-internet/introduction. NERA Economic Consulting.

Gautier, A. and Somogyi, R. (2020). Prioritization vs zero-rating: Discrimination on the internet. *International Journal of Industrial Organization*, 73:102662.

Hoernig, S. and Monteiro, F. (2020). Zero-rating and network effects. *Economics Letters*, 186:108813.

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. (2021). Zero-rating and vertical content foreclosure. *Information Economics and Policy*, 55:100899.

Jullien, B. and Sand-Zantman, W. (2018). Internet regulation, two-sided pricing, and sponsored data. *International Journal of Industrial Organization*, 58:31–62.

Lorenzon, E. (2022). Zero-rating, content quality, and network capacity. *Information Economics and Policy*, 58:100965.

Mussa, M. and Rosen, S. (1978). Monopoly and product quality. *Journal of Economic Theory*, 18(2):301–317.

OECD (2024). Information and communication technology. https://www.oecd-ilibrary.org/science-and-technology/information-and-communication-technology-ict/indicator-group/english_04df17c2-en. Accessed: 16 January 2024.

Preta, A. and Peng, P. (2020). Discrimination and neutrality on the internet: the zero rating case. Unpublished.

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 184965, International Telecommunications Society (ITS).

Somogyi, R. (2016). The Economics of Zero-Rating and Net Neutrality. LIDAM Discussion Papers CORE 2016047, Université catholique de Louvain, Center for Operations Research and Econometrics.

Appendix A. Reproducibility of computational results for the *ISP*' maximization problem.

To enable reproducibility of our results, we define a grid of uniformly spaced parameters combinations. The grid bounds and the inter-spacing for each parameter are shown in Table A.1. The bounds were chosen following two rules: if the parameter has natural bounds, we choose these bounds, for example, parameters γ and θ lay between 0 and 1; we also have that $1 = \alpha_2 < \alpha_1 \leq 2$; c and d^T have 0 as their lower bound; and d^S has d^T as its lower bound. Otherwise, bounds as set to guarantee *interior* and *valid* solutions. These latter requirements are carefully explained in the main text.

Table A.1: Range of the market parameter values and incremental size

Parameter	Lower bound	Upper bound	Step
α_1	1.10	2.00	0.10
γ	0.00	0.98	0.02
θ	0.02	0.98	0.04
d^T	0.10	1.00	0.10
d^S	d^T+0.10	1.00	0.10
c	0.01	1.00	0.01

The combination of the six parameters (recall that α_2 is fixed to 1) leads us to compute 56,250,000 different instances of the ISP's maximization problems. Numerical methods are used to find the optimal menu of two-part tariffs for each instance. Not all instances find an interior and valid solution, so that under neutral network the number of solutions reached is 1,221,412.

Appendix B. Neutral network

Appendix B.1. A single two-part tariff

Suppose now that the ISP maximizes its profit function by choosing a network capacity μ and a single two-part tariff (H, τ), i.e., $H = H^T = H^S$ and $\tau = \tau^T = \tau^S$. The ISP' s profit maximization is similar to the one above, but only considering the participation constraints. To better compare the findings, we assume that the ISP will set the common two-part tariff such that it guarantees that both consumer types subscribe with Internet.

By backward induction, in the third-period each consumer maximizes her own utility function, given the ISP and the CP's equilibrium strategies:

$$\begin{aligned} Max_{q_1^m, q_2^m} \quad & U^m(q_1^m, q_2^m | p_1, p_2, H, \tau, \mu) \\ s.t. \quad & q_1^m, q_2^m > 0 \end{aligned}$$

The first order conditions give us the equilibrium consumption function that each type of consumer makes of each of the content providers:

$$\begin{aligned} q_i^T &= \frac{1}{1-\gamma^2}\Big[\alpha_i - \gamma\alpha_j - (p_i - \gamma p_j) - (1-\gamma)(\theta d^T + \tau)\Big], \\ q_i^S &= \frac{1}{1-\gamma^2}\Big[\alpha_i - \gamma\alpha_j - (p_i - \gamma p_j) - (1-\gamma)((1-\theta)d^S + \tau)\Big]. \end{aligned}$$

It is easy to show that T-consumers' equilibrium consumption of any CP will be higher than that of the S-consumers if and only if $\theta d^T < (1-\theta)d^S$. As $d^T < d^S$, this means than only when the proportion of T-consumers is high enough, the S-consumers' equilibrium consumption will exceed that of T-consumers.

In the second period, the content providers' equilibrium strategies are to set simultaneously and independently the prices that maximize their profit functions, given the ISP equilibrium strategy, and the equilibrium quantities in the continuation game:

$$\begin{aligned} Max_{p_j} \pi_j^{CP}(p_j \mid \tau) &= p_j[\theta q_j^T(p_1, p_2, \tau) + (1-\theta) q_j^S(p_1, p_2, \tau)] \\ s.t. \quad & p_j > 0. \end{aligned}$$

By subgame perfection, substituting $q_j^T(p_1, p_2, \tau)$ and $q_j^S(p_1, p_2, \tau)$ by the equilibrium functions of the third period, and solving,

$$\begin{aligned}
p_1 &= \frac{1}{4-\gamma^2}\big[(2-\gamma^2)\alpha_1 - \gamma\alpha_2 - (1-\gamma)(2+\gamma)(\bar{d}+\tau)\big], \\
p_2 &= \frac{1}{4-\gamma^2}\big[(2-\gamma^2)\alpha_2 - \gamma\alpha_1 - (1-\gamma)(2+\gamma)(\bar{d}+\tau)\big].
\end{aligned}$$

where $\bar{d} = \theta^2 d^T + (1-\theta)^2 d^S$. Then, we see the CP's quality differentiation plays a role in the equilibrium price setting. In fact, the difference of CP's equilibrium content prices is $p_1 - p_2 = \frac{1+\gamma}{2+\gamma}(\alpha_1 - \alpha_2)$, therefore it is always the case that $p_1 > p_2$

By subgame perfection, we substitute equilibrium prices into equilibrium consumption. Then, the equilibrium consumption only depends on the unit overage charge τ. Therefore, in the first period, the ISP's equilibrium strategy is to set its capacity and its two-part tariff, which are also an equilibrium in the continuation game. This implies the maximization of,

$$\begin{aligned}
Max_{H,\tau,\mu}\pi^{ISP}(H,\tau,\mu) &= H + \tau[\theta(q_1^T(\tau) + q_2^T(\tau)) + (1-\theta)(q_1^S(\tau) + q_2^S(\tau))] - \\
&\quad c(\mu + \frac{1}{2}\mu^2) \\
s.t. &\quad U^T(q_1^T(\tau), q_2^T(\tau)) \geq 0 \\
&\quad U^S(q_1^S(\tau), q_2^S(\tau)) \geq 0 \\
&\quad H, \tau, \mu \geq 0.
\end{aligned}$$

At equilibrium, the ISP will extract the entire surplus of at least one of the consumer types. Otherwise, the ISP would have room to increase the hookup fee to get higher profit. Let us say that in the optimal solution the m-consumer has a utility equal to zero, then it can be algebraically shown that the ISP's equilibrium strategy is,

$$\begin{aligned}
\tau^m &= \frac{1}{6-4\gamma}[(1-\gamma)(\alpha_1+\alpha_2) + 2(2-\gamma)d^m] - \bar{d} \\
\mu^m &= \frac{d^m}{c} - 1
\end{aligned}$$

and H^m is obtained from $U^m(q_1^m(\tau^m), q_2^m(\tau^m)) = 0$.

In addition, at equilibrium, for a wide range of θ, the ISP fully extracts the surplus of the T-consumers. Only when the market consists almost exclusively of S-type consumers (θ has a very low value), the ISP finds it optimal to fully extract the surplus of the S-consumer type.

Appendix B.2. Equilibrium consumption under a single two-part-tariff as compared to a menu of tariffs

Under a two-part tariff menu, the ISP keeps constant the capacity investment but designs a mechanism for traffic control, considering that a potential increase in consumption will reduce the network flow speed, which, in turn, may reduce content consumption. The mechanism works as follows. On the one hand, it increases the ISP's profits by charging different pairs (H, τ) to consumers. In this way, it charges a lower τ to the less traffic speed demanding consumers and increases τ for the more demanding ones. The precise explanations is as follows.

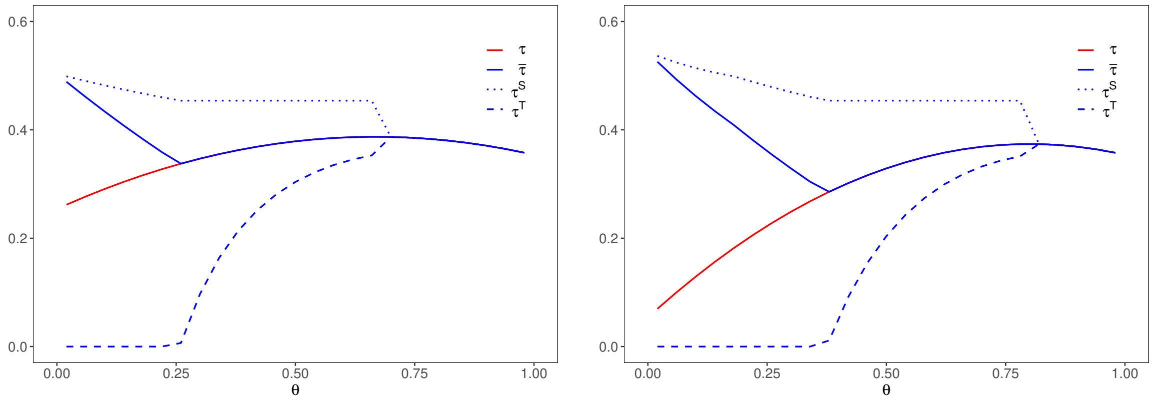
Under a single two-part tariff, say (H, τ), whenever $\theta d^T < (1-\theta)d^S$, the T-consumer equilibrium quantities of both CPs is always higher than that of S-consumers. Therefore, depending on θ the ISP could design a profitable two-part tariff menu which could alleviate the congestion externally. Thus, in two-part tariff menu, for low θ the ISP fixes $\tau^T = 0$ and $\tau^S > 0$, with $\tau^S > \tau > \tau^T = 0$. In this way, q_i^T content consumption will increase and q_i^S consumption will decrease, respectively, with respect the single two-part tariff, alleviating the S-consumer congestion externality. Moreover, since the CP's prices at two-part tariff menu depend negatively on $\bar{\tau}$ (recall $\bar{\tau} = \theta\tau^T + (1-\theta)\tau^S$) and $\bar{\tau} > \tau$, then both CP's prices, p_1 and p_2, will decrease with respect to those under a single two-part tariff. However, since θ is quite small, the (weighted) aggregate total consumption is such that $Q_1^{TM} + Q_2^{TM} < Q_1^{ST} + Q_2^{ST}$, where superscript TM means Tariff menu and ST means single tariff. With this, not only the S-externality is alleviated but also there is an increase in the network flow, with permits more consumption by T-consumers. Obviously, both CP's profits and U^S will be smaller but the ISP's profits will be higher because of higher H^T and H^S. The utility of T-consumer is kept equal to zero as under a single two-part tariff.

For intermediate θ's, but still with $\theta d^T < (1-\theta)d^S$ the ISP can only profitably design new τ's such that: i) $\tau^S > \tau = \bar{\tau} > \tau^T > 0$; and ii) $\theta d^T + \tau^T < (1-\theta)d^S + \tau^S$. Given i), the CP's new prices, p_1 and p_2 under the two-part tariff menu are the same as under a singe two-part tariff scheme. Consequently, $Q_1^{TM} + Q_2^{TM} = Q_1^{ST} + Q_2^{ST}$. However, by ii) q_i^T quantities will increase and q_i^S consumptions will decrease, alleviating the S-consumer congestion externality, but the network flow is not altered. Both CP's profits are the same than under the single two-part tariff scheme and U^S will be smaller but the ISP's profits will be higher because of higher H^T and H^S. The utility of T-consumer is kept equal to zero as under a single two-part tariff.

Finally, for higher θ such that $\theta d^T \geq (1-\theta)d^S$, the S-consumers consumes more from each CP than T-consumers do under a single two-part tariff. The tariff menu is such that $\tau^S = \tau = \bar{\tau}^S = \tau^T$, and the CP's new prices, p_1 and p_2 under the two-part tariff menu are the same as under a single two-part tariff scheme and then $Q_1^{TM} + Q_2^{TM} = Q_1^{ST} + Q_2^{ST}$. There is a θ such that the optimal tariff that the ISP can fix makes $\theta d^T + \tau^T = (1-\theta)d^S + \tau^S$ and then, the T-consumer consumption is the same than that of S-consumers under the two-part tariff menu. From this θ on, the ISP does not have any room to profitably redirect the traffic flow and hence it comes back to a single two-part tariff. (see Figure B.11, which illustrates the different $\tau's$ choices).

The result is that the (weighted) aggregate consumption under a tariff menu is smaller than or equal to the one under a single two-part tariff (see top row in Figure B.12). Thus, aggregate consumption does not increase with a tariff menu, but changes its allocation between the two types of consumers. Specifically, when the proportion of the less speed demanding consumers (T-consumers) is small, the (weighted) aggregate total consumption under a tariff menu is smaller than the one under a single two-part tariff. As the proportion of T-consumers increases, their total consumption increases, and from some proportion on, they consume more than S-consumers. In general, as the bottom-row in Figure B.12 reveals, under a tariff-menu T-consumers (S-consumers) consume more (less) than under a single two-part tariff. If the proportion of the less demanding is quite high, each type of consumers consumes the same under the two tariff regimes.

Figure B.11: Equilibrium overage charges (τ) under single tariff and menu of tariffs.



(a) $d^S = 0.2$ (Low consumer differentiation)

(b) $d^S = 0.5$ (High consumer differentiation)

Notes: Both figures show: the equilibrium overage charge under a single tariff τ; the equilibrium overage charges under a menu of tariffs τ^T and τ^S; and the average of the equilibrium overage charges under a menu of tariffs $\bar{\tau}$. Consumption has been computed for parameters $\alpha_1 = 1.5$, $\alpha_2 = 1$, $\gamma = 0.2$, $d^T = 0.1$ and $c = 0.01$, and scale has been set equal for plots of the same row.

Appendix C. Restricted network: Zero-rating

Appendix C.1. ISP zero-rates CP_1

Let $\bar{d} = \theta^2 d^T + (1-\theta)^2 d^S$ and $\bar{\tau} = \theta\tau^T + (1-\theta)\tau^S$.

Third-period: The consumers' equilibrium strategies. The first order conditions give us the equilibrium consumption functions that each type of consumer makes of each of the content providers, for $i, j = 1, 2$, $i \neq j$,

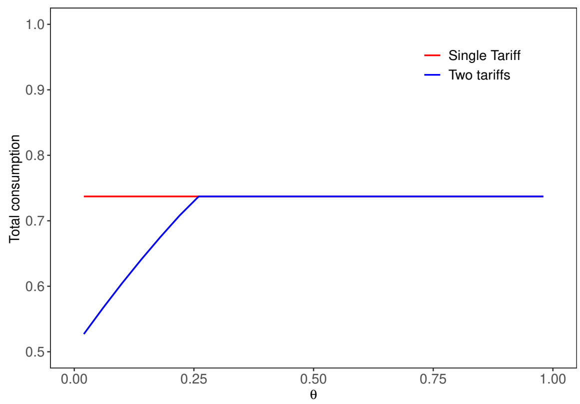
$$q_1^T = \frac{1}{1-\gamma^2}\Big[\alpha_1 - \gamma\alpha_2 - (p_1 - \gamma p_2) - (1-\gamma)\theta d^T + \gamma\tau^T\Big] \tag{C.1}$$

$$q_1^S = \frac{1}{1-\gamma^2}\Big[\alpha_1 - \gamma\alpha_2 - (p_1 - \gamma p_2) - (1-\gamma)(1-\theta) d^S + \gamma\tau^S\Big] \tag{C.2}$$

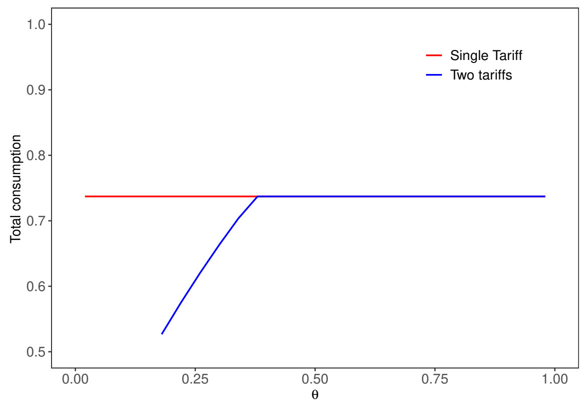
$$q_2^T = \frac{1}{1-\gamma^2}\Big[\alpha_2 - \gamma\alpha_1 - (p_2 - \gamma p_1) - (1-\gamma)\theta d^T - \tau^T\Big] \tag{C.3}$$

$$q_2^S = \frac{1}{1-\gamma^2}\Big[\alpha_2 - \gamma\alpha_1 - (p_2 - \gamma p_1) - (1-\gamma)(1-\theta) d^S - \tau^S\Big] \tag{C.4}$$

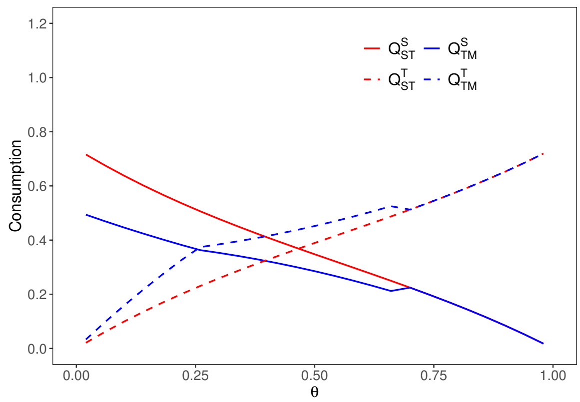
Figure B.12: Equilibrium aggregate consumption (top row) and equilibrium consumption by consumers' types (bottom row).



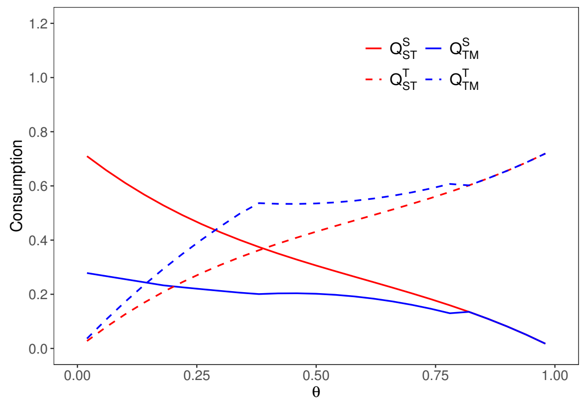
(a) $d^S = 0.2$ (Low consumer differentiation)



(b) $d^S = 0.5$ (High consumer differentiation)



(c) $d^S = 0.2$ (Low consumer differentiation)



(d) $d^S = 0.5$ (High consumer differentiation)

Notes: Top row displays equilibrium weighted aggregate consumption $Q = \theta Q^T + (1-\theta)Q^S$, with $Q^m = q_1^m + q_2^m$, $m = T, S.$, under a single tariff and a tariff menu. In the bottom row sub-index ST corresponds to single tariff and TM to tariff menu. Consumption has been computed for parameters $\alpha_1 = 1.5$, $\alpha_2 = 1$, $\gamma = 0.2$, $d^T = 0.1$ and $c = 0.01$, and scale has been set equal for plots of the same row.

Second period: Content Providers' equilibrium strategies

$$
\begin{aligned}
p_1 &= \frac{1}{4-\gamma^2}\big[(2-\gamma^2)\alpha_1 - \gamma\alpha_2 - (1-\gamma)(2+\gamma)\bar{d} + \gamma\bar{\tau}\big] && \text{(C.5)}\\
p_2 &= \frac{1}{4-\gamma^2}\big[(2-\gamma^2)\alpha_2 - \gamma\alpha_1 - (1-\gamma)(2+\gamma)\bar{d} - (2-\gamma^2)\bar{\tau}\big] && \text{(C.6)}
\end{aligned}
$$

Appendix C.2. ISP zero-rates CP_2

Third-period: The consumers' equilibrium strategies.

$$
\begin{aligned}
q_1^T &= \frac{1}{1-\gamma^2}\Big[\alpha_1 - \gamma\alpha_2 - (p_1 - \gamma p_2) - (1-\gamma)\theta d^T - \tau^T\Big] && \text{(C.7)}\\
q_1^S &= \frac{1}{1-\gamma^2}\Big[\alpha_1 - \gamma\alpha_2 - (p_1 - \gamma p_2) - (1-\gamma)(1-\theta) d^S - \tau^S\Big] && \text{(C.8)}\\
q_2^T &= \frac{1}{1-\gamma^2}\Big[\alpha_2 - \gamma\alpha_1 - (p_2 - \gamma p_1) - (1-\gamma)\theta d^T + \gamma\tau^T\Big] && \text{(C.9)}\\
q_2^S &= \frac{1}{1-\gamma^2}\Big[\alpha_2 - \gamma\alpha_1 - (p_2 - \gamma p_1) - (1-\gamma)(1-\theta) d^S + \gamma\tau^S\Big] && \text{(C.10)}
\end{aligned}
$$

Second period: Content Providers' equilibrium strategies

$$
\begin{aligned}
p_1 &= \frac{1}{4-\gamma^2}\big[(2-\gamma^2)\alpha_1 - \gamma\alpha_2 - (1-\gamma)(2+\gamma)\bar{d} - (2-\gamma^2)\bar{\tau}\big] && \text{(C.11)}\\
p_2 &= \frac{1}{4-\gamma^2}\big[(2-\gamma^2)\alpha_2 - \gamma\alpha_1 - (1-\gamma)(2+\gamma)\bar{d} + \gamma\bar{\tau}\big] && \text{(C.12)}
\end{aligned}
$$