International Journal of Industrial Organization 73 (2020) 102662

Contents lists available at ScienceDirect

International Journal of Industrial Organization

journal homepage: www.elsevier.com/locate/ijio

Prioritization vs zero-rating: Discrimination on the internet☆

Axel Gautier[a,b,e,*], Robert Somogyi[c,d]

[a] HEC Liege, University Liege, LCII. Bat B31, Quartier Agora, Place des orateurs 3, Liege B4000, Belgium
[b] CORE (UCLouvain), Belgium
[c] Department of Finance, Budapest University of Technology and Economics, Hungary
[d] Institute of Economics, Centre for Economic and Regional Studies, Hungary
[e] CESifo (Munich) Germany

ARTICLE INFO

Article history:
Received 12 June 2019
Revised 9 August 2020
Accepted 7 September 2020
Available online 17 September 2020

JEL classification:
D21
L12
L51
L96,

Keywords:
Net Neutrality
Paid Prioritization
Zero-rating
Sponsored Data
Data Cap
Congestion

ABSTRACT

This paper analyzes two business practices on the mobile internet market, paid prioritization and zero-rating. These practices allow the internet service provider to discriminate different content types. With prioritization, the ISP delivers content at different speeds; with zero-rating, the ISP charges different prices. In recent years these practices have attracted considerable media attention and regulatory interest. When the asymmetry between content providers is limited, in particular with regard to their ability to attract traffic or to monetize it, we first show that the ISP can extract more surplus from consumers by privileging the relatively weaker content and restoring symmetry between content providers. Next, we show that the ISP chooses prioritization when traffic is highly valuable for content providers and congestion is severe, and zero-rating in all other cases. Finally, we find that a policy banning prioritization can lead to zero-rating and a reduction in consumer surplus.

1. Introduction

Net neutrality is the principle of equal treatment of all data packages sent over the internet, irrespectively of their content, origin, destination and type of equipment used to access it. This regulatory principle prohibits discrimination on the internet and mandates Internet Service Providers (ISPs) to treat all data packets equally. In recent years, two business practices, prioritization and zero-rating, have been widely discussed and debated both in policy circles and the media. These practices violate the net neutrality principle and allow the ISP to discriminate between content in terms of *quality* for prioritization and in terms of *price* for zero-rating. Thus they allow the ISP to add an additional layer of differentiation between

☆ We would like to thank Paul Belleflamme, Marc Bourreau, Anna D'Annunzio, Johannes Johnen, Bruno Jullien, Sergey Lychagin, Norbert Maier and conference and seminar participants at the 12th Digital Economics Conference (Toulouse), VIII IBEO Media Economics workshop (Alghero), 4th Industrial Organization in the Digital Economy workshop (Liege), 10th Digital Economics conference (Paris), EEA (Cologne), EARIE (Athens), Institute of Economics, CERS (Budapest) and CORE (Louvain-la-Neuve) for discussion and comments. This research was funded through the ARC grant for Concerted Research Actions, financed by the French speaking Community of Belgium. Somogyi has been a Carrier Integration Fellow of the CERGE-EI Foundation while undertaking the research in this paper. The research reported in this paper was supported by the BME NC TKP2020 grant of NKFIH Hungary.

* Corresponding author.
E-mail addresses: agautier@uliege.be (A. Gautier), somogyi@finance.bme.hu (R. Somogyi).

https://doi.org/10.1016/j.ijindorg.2020.102662

Content Providers (CPs) beyond their intrinsic differentiation. Our objective in this paper is to build a model that evaluates the costs and benefits of such an empowerment of the ISP and to contrast and compare the two practices.

The practice of *prioritization* consists of creating a fast lane on the internet where privileged content circulates in case of congestion. The prioritized content has a better quality –faster delivery– while the non-prioritized content types experience congestion in the form of throttling, jitter and/or delays. Prioritization is thus equivalent to quality differentiation. The *zero-rating* practice consists of creating financial differentiation between content providers. In this case, the ISP makes the consumption of some data packets more expensive than others. A typical example of financial discrimination is mobile data plans with zero-rated content. With such a contract, users subscribe to a package with a monthly data cap but the usage of some content (e.g. Facebook Messenger or Netflix) does not count against this data cap, i.e. the ISP charges different marginal rates for different types of content. Zero-rating is thus equivalent to price discrimination. With prioritization and zero-rating, the ISP directly influences the competition between content providers. Therefore, they might be ready to compensate the ISP for having priority or being zero-rated.

Net neutrality is the only regulation that is specific to the internet and it is a highly contentious issue (see for instance the detailed discussion in Greenstein et al., 2016). Different countries adopt widely different approaches to net neutrality and regulators often have a contrasting attitude towards prioritization and zero-rating. The objective of this paper is to compare the two types of business practices within the same model, thus evaluating the often differentiated policy towards them.

We consider a monopolistic ISP that connects CPs and consumers, i.e. the ISP is a two-sided intermediary. In a neutral internet, the ISP must charge the same price to all consumers and CPs, the latter being equal to zero.[1] In a non-neutral internet, the ISP can discriminate between content, and our model shows that it does so in equilibrium. Discrimination gives an advantage on the content market to the privileged CP, an advantage for which the ISP may ask for compensation. This, as a consequence, changes the price structure of the ISP, which then charges a positive price on the CP side of the market as well.

Our comparison will focus on the ability of the ISP to differentiate content and the impact this empowerment has on prices, consumer surplus and profits. More specifically, we construct a model to compare three regimes: prioritization (P), zero-rating (ZR) and net neutrality (NN). In our model, discrimination has no value as such, instead, it is considered as a tool for the ISP to alter competition on the content market.[2] In our model, there are two competing content providers. The CPs are financed by advertising, and they compete to attract internet users. They offer differentiated content, represented by the standard Hotelling line. We suppose asymmetric content providers with one CP (called the strong CP) benefiting from a larger home market, i.e. more consumers have an intrinsic preference for its content. Importantly, we assume that the asymmetry between content providers is limited in several aspects, i.e. in the size of the home market, in the advertising revenue they can generate and also in their potential reach of new consumers in the model variants where demand is variable. To access content, users must subscribe to an internet service provider (ISP), that we suppose to be in a monopolistic position. The ISP faces a capacity constraint and thus may not be able to deliver the best available quality to all users.[3]

Under prioritization, the ISP discriminates between content providers and gives the content supplied by CP_i priority access over the content supplied by CP_j. Therefore, if there is congestion on the internet, the download quality is not identical for the two types of content. The priority content provider CP_i may or may not compensate the ISP for getting priority, but consumers pay the same price to the ISP for accessing both types of content.[4] Under zero-rating, consumers are not charged the same price by the ISP when they access the content of CP_i as when they access the content of CP_j. In other words, the ISP financially discriminates the two content providers, but they remain identical in terms of download quality. The zero-rated content provider CP_i may or may not compensate the ISP for this service.[5] Under net neutrality, the ISP cannot discriminate between content providers, so this regime can serve as a natural benchmark. In particular, net neutrality prohibits quality differentiation like prioritized content and financial differentiation like applying a lower rate for some types of content. A consequence is the absence of financial transfers between the ISP and the CPs under net neutrality.

Both prioritization and zero-rating create distortions on the CP market compared to net neutrality, i.e. market shares on the CP market are affected by the ISP. The ability to distort content consumption depends, in the case of prioritization, on the bandwidth capacity with a lower bandwidth creating more distortion and, in case of zero-rating, on the consumers' preferences for the strong content with a higher share of captive consumers creating more distortion.

[1] In other words, we use a strict definition of net neutrality, interpreted as a price regulation fixing a zero price on one side of the market, the CP side, see, e.g. Economides and Tåg (2012).

[2] There are other costs and benefits associated with the two practices that are not considered in our model. For instance, prioritization is considered to be an efficient tool to manage congested data traffic when content types have different sensitivity to delay (Peitz and Schuett, 2016).

[3] The coronavirus epidemic has revealed the importance of capacity constraints even in fixed broadband networks. For example, in New York median download speeds reportedly dropped 24% when the lockdown started, see The New York Times https://tinyurl.com/vdu6ysh.

[4] We use the term "prioritization" instead of "paid prioritization" (the standard way to refer to this type of quality differentiation) because we allow the ISP to prioritize a CP even in the absence of transfers between the ISP and the CP.

[5] Some use the term "sponsored data" to describe this business practice. However, "sponsored" implies a transfer from the CP to the ISP which does not necessarily occur in our model, nor in real life. In the US, most mobile carriers zero-rate some video content. The Binge On program of T-Mobile is a free zero-rating program, and all content providers that meet some technical requirements can apply freely. On the contrary, to be admitted in the zero-rating plan of AT&T (called Sponsored Data), content providers have to compensate the ISP.

Our first main result is that with price discrimination (zero-rating), the ISP can extract more surplus from consumers than with quality discrimination (prioritization). Hence, in the absence of revenues from CPs, the ISP prefers zero-rating to prioritization and net neutrality. Moreover, we show that the ISP chooses prioritization only when traffic is highly valuable for CPs and congestion is severe. In all other cases, i.e. for low congestion and/or low value of traffic, the ISP chooses to use zero-rating. The intuition is the following. Without payments from the CPs, the ISP prefers to discriminate in price rather than in quality, i.e. it prefers zero-rating to prioritization. Then, the ISP will choose prioritization only if it receives a large payment from the prioritized content provider. This will be the case when traffic has much value for content providers *and* prioritization is associated with substantial benefits on the market for content, i.e. when it creates large distortions in content consumption, which is the case when congestion is severe. In all other cases, the ISP chooses zero-rating.

When prioritization is implemented, it can benefit both the ISP and the consumers. The reason is that the ISP has less ability to extract surplus from consumers under prioritization than under zero-rating, but it can compensate this by larger payments from the content provider side. Thus for severe congestion, the price paid by consumers under prioritization will be low, while the contribution of the CPs will be high and there is a form of 'cross-subsidization' of consumers by the CPs. Such a situation can arise as the ISP is a two-sided platform and a higher price on one side (the CP side) can lead to a lower price on the other side (the consumer side). Hence, consumers might be better-off under prioritization than under zero-rating and, in some cases than under net neutrality. We show that the interests of the consumers and the ISP could be congruent on the choice of prioritization, whereas it is never the case for zero-rating.

Our second main result is to show that both under zero-rating and prioritization, the ISP will privilege the weak content provider.[6] Discrimination against the strong content is a way for the ISP to mitigate the initial asymmetry between CPs and increase the revenue it collects from consumers by making them pay relatively more to consume the strong content. Hence, the ISP prefers to privilege the weaker content provider. Furthermore, in a fixed demand configuration, prioritization and zero-rating only displace content consumption and the market share lost by one CP is gained by the other one. Therefore the two CPs will bid the same amount to be the privileged content if they value traffic equally.

Next, we relax the limited asymmetry assumption between content types and we consider cases in which the strong content provider can attract more consumers, more traffic or more advertising revenue than the weak one. Intuitively, when the strong content has a large advantage either in the size of its home market, or in its reach of new potential consumers, or in its revenue-generating ability; zero-rating the strong CP or giving priority to it may generate more revenue from both the user side and the CP side. We show that this asymmetry must be large enough in order to convince the ISP to give a preferential treatment to the strong instead of the weak content. In other words, we derive upper limits on asymmetry between content types.

Finally, we consider several further extensions of the model. First, we consider the possibility for the ISP to combine zero-rating and prioritization. In this case, priority will be, as in the baseline model, given to the weak CP, but we show that zero-rating reinforces prioritization only when congestion is limited. Otherwise, it is optimal to prioritize the weak content and zero-rate the strong one. In the second extension, we consider the incentives of the ISP to invest in network improvements, and we show that, under prioritization, incentives to invest are suboptimal. The intuition is that the more severe congestion is, the more impact prioritization has on the market for content and hence, the more surplus the ISP can extract from content providers. For this reason, incentives to invest in capacity extension are limited. On the contrary, we find that under zero-rating, the ISP's investment incentives are socially optimal. Third, we discuss why we believe that our results considering a monopolistic ISP apply to some settings with ISP competition as well.

Our research project is closely related to the rich body of literature in theoretical industrial organization about net neutrality. This literature has focused on four practices that are violations of the net neutrality principle: exclusion of lawful content (Broos and Gautier, 2017), throttling, prioritization and zero-rating.[7]

Concerning prioritization, in the context of congested networks, a two-tiered internet service with a fast lane for prioritized content can be considered as a more efficient tool for traffic management than a strict net neutrality rule, see Peitz and Schuett (2016) for an explicit analysis of this issue. Another important aspect of prioritization is that it changes the market for internet content, which in turn affects the optimal prices ISPs charge to both sides of the market.[8] This aspect of prioritization is also made explicit in our modeling of prioritization (and zero-rating as well). Finally, departing from a neutral internet and pricing congestion also changes the incentives of the ISP to invest in network capacity. There is no consensus in the literature on the impact of prioritization on the investments in capacity. In our model, prioritization reduces the incentives of the ISP to upgrade the network as congestion is then a valuable resource.

Our modeling of prioritization is closest to Choi and Kim (2010) and Cheng et al. (2011) as we also use CPs located on a Hotelling segment populated with single-homing consumers. The logic of these models is also similar: prioritization creates value by distorting the content market, and the ISP is able to capture part of the increased value from the prioritized CP.

[6] In the US, AT&T zero-rates its affiliated video streaming subsidiary DirectTV, and Verizon also zero-rates is own subsidiary, go90. These services are arguably weaker than Youtube or Netflix. Although we model the ISP as a monopoly, we discuss in Section 8 the reasons why we believe our results apply to some competitive ISP settings as well.

[7] See Easley et al. (2018) for a detailed description of network management practices that violate net neutrality (and those that do not).

[8] Important contributions include Choi and Kim (2010), Cheng et al. (2011), Economides and Hermalin (2012), Economides and Tåg (2012), Krämer and Wiewiorra (2012), Bourreau et al. (2015), Reggiani and Valletti (2016) and Choi et al. (2018). For some relatively recent surveys, see Greenstein et al. (2016) or Krämer et al. (2013).

Fig. 1. Content providers on the Hotelling line.

Some of our results are in line with previous findings. Choi and Kim (2010, proposition 1), finds that for high levels of advertising revenue the ISP prefers prioritization to net neutrality. In our model, prioritization is also the ISP's preferred regime for high advertising revenue whenever it is combined with a relatively high congestion level. Otherwise, the preferred regime is zero-rating and never net neutrality. Clearly, this result has welfare effects with a potential to alter policy implications. Our main contribution is considering zero-rating agreements as an alternative business strategy to prioritization for the ISP to create distortion on the content market.

To date there have been few formal economic studies modeling zero-rating. Jullien and Sand-Zantman (2018) model zero-rating as a coupon from content providers to end users, a potential way to overcome the misallocation problem stemming from the free content business model. Somogyi (2018) investigates the trade-off between congestion and increased utility from consumption by modeling data caps explicitly, distinguishing the two types of zero-rating programs (open and exclusionary) currently in use. Jeitschko et al. (2017) focus on vertical integration between a content provider and the internet service provider in a zero-rating regime. Inceoglu and Liu (2019) focus on the consumer side of the internet market and model zero-rating as a screening device for price discrimination. Hoernig and Monteiro (2018) analyze zero-rating when payments from content providers to the ISP are banned and the profitability of zero-rating arises solely from network effects. Cho et al. (2016) model zero-rating with two CPs that are located at the extremes of a Hotelling line, and highlight the importance of the transportation cost parameter in determining the equilibrium market outcomes. Schnurr and Wiewiorra (2018) model exclusionary zero-rating offers with and without side-payments from the CPs to the ISP, assuming tariffs with data caps and content-specific prices; and compare the outcomes to conventional, i.e. one-sided pricing practices. Finally, Krämer and Peitz (2018) highlight the prevalence of throttling of zero-rated content. Our model is well suited to analyze this case, as throttling of a content provider can be seen as prioritizing its rivals.

The fact that congestion is still a salient feature of the internet today has recently been documented by the thorough empirical studies of Nevo et al. (2016) and Malone et al. (2017). In particular, by analyzing fixed residential broadband data they estimate a large willingness-to-pay to avoid congestion. Several factors, including hard technological constraints make mobile internet even more susceptible to congestion, including the scarcity of spectrum and limits on cell size reduction (see e.g. Clarke, 2014). Indeed, Heikkinen and Berger (2012) find that in most countries packet loss, a standard measure of congestion, is higher on the mobile internet.

Finally, our model is also related to the literature of simultaneous horizontal and vertical product differentiation. The two seminal articles in the literature; Economides (1989), and Neven and Thisse (1989) show that firms choose maximal differentiation in one dimension and minimal differentiation in the other dimension. One finding of our model is similar in spirit to these results. In particular, when prioritization and zero-rating are both available tools for the ISP to differentiate CPs it finds it optimal to use only one of them, zero-rating.

2. A model of zero-rating and prioritization

In the model, a monopolistic internet service provider (ISP) acts as an intermediary between competing content providers (CP) and consumers.

2.1. Content providers

There are two contents providers and we use a Hotelling line of length 1 with linear transportation cost τ to model horizontal differentiation between content types. CPs are asymmetric, CP_S is located at point $a \geq 0$ on the unit interval and CP_W is located at point 1. CP_S, hereafter the strong content provider, is in a privileged position with a larger natural market than its rival CP_W, hereafter the weak content provider (see Fig. 1).[9]

The strong CP can create such an asymmetry in a number of ways, including a larger installed base thanks to a first-mover advantage, or simply providing more attractive content. Without this initial asymmetry between content, the ISP has no reason to financially differentiate the content types. The parameter a cannot be too large, otherwise the consumer located at the left extreme of the interval will stop consuming and the market will not be covered anymore. The following assumption limits the asymmetry between content types and guarantees a covered market in all the market configurations we deal with.

Assumption 1. $a < \overline{a} \equiv \min\{\frac{2}{7};\ \frac{1}{4\tau}\}$.

[9] This formulation is similar to the Hotelling model with hinterlands (Inderst and Peitz (2014)) in which one firm has a mass of consumers (a) with a strong preference for its content. Moreover, from the formulation of $U(x)$ below, it is clear that our setting is equivalent to a model of vertical differentiation where both firms are located at the extremities and all consumers have a valuation $v_S = v + a\tau$ for the strong content and $v_W = v$ for the weak content.

In a covered market, the aggregate demand of content is fixed and we will relax this assumption in Section 7.1.

Consumers who are connected to the internet single-home, that is they choose to consume the content of either CP_S or CP_W. Content is offered for free by the CPs and the content provision service is financed by advertising revenues.[10] We suppose that the CPs collect an exogenous advertising revenue ρ per viewer. Content providers' operating costs are normalized to zero. If we denote the total mass of viewers of content $i \in \{S, W\}$ by n_i, the profits of the content providers are $\pi_{CP_i} = \rho n_i$.

2.2. Internet service provider

There is a single ISP. Consumers need to be connected to the ISP to access the content offered by the CPs. Each connected consumer will consume one unit of content either at CP_S or at CP_W and the total demand of the ISP is $n_S + n_W$.

The ISP has a transmission capacity of κ. Limited capacity may lead to congestion if the bandwidth is insufficient to carry all the traffic. The consumption of time-sensitive content will be altered in case of congestion and consumers will experience jitter, delays, interruptions or a degradation of the content quality (throttling). We represent the surfing experience quality by a parameter q, that can be interpreted as the probability of on-time delivery (Peitz and Schuett, 2016). This probability is equal to the ratio of the bandwidth to the traffic: $q = \min[1, \frac{\kappa}{n_S+n_W}]$. If $n_S + n_W < \kappa$, there is no congestion on the internet and all content can be delivered on time: $q = 1$. If $n_S + n_W > \kappa$, the internet is congested and the ISP can no longer deliver the quality $q = 1$ to all users. If there is no discrimination between content providers, the probability of on-time delivery is the same for all users and all content and equals $q = \frac{\kappa}{n_S+n_W} < 1$. We will assume that capacity is scarce, otherwise prioritization is meaningless. More precisely, we assume that the highest quality cannot be delivered to all consumers when the market is covered.

Assumption 2. $\kappa < 1$.

In the prioritization regime, the ISP will give priority to content i over content j and the probability of on-time delivery will be higher for content i than for content j: $q_i \geq q_j$. If $n_i < \kappa$, part of the bandwidth will be used for the prioritized content that will be delivered on time: $q_i = 1$ while the remaining bandwidth $(\kappa - n_i)$ will be used for the non-priority content that will experience delays: $q_j = \frac{\kappa - n_i}{n_j} < 1$. If $n_i > \kappa$, even the priority content will experience delays $(q_i = \frac{\kappa}{n_i} < 1)$ and the non-priority content has the lowest possible quality $(q_j = 0)$. Note that consumers may still consume content j even if q_j is equal to zero.

While this rationing scheme is very tractable, it also shares some valuable properties with the M/M/1 queuing system, which is believed to be a good approximation of real rationing processes and is used by e.g. Choi and Kim (2010), Cheng et al. (2011), Choi et al. (2018).[11] First, the average quality in our model is the same (κ) under prioritization and net neutrality, analogously to the average waiting time being independent of regimes in the above models. Moreover, in interior solutions (which turn out to be the only equilibrium configuration) the quality differential decreases in capacity κ: $\frac{\partial(q_i - q_j)}{\partial\kappa} \leq 0$, just like the difference in waiting time is reduced in M/M/1 queues.

Consumers pay the ISP for accessing the internet. Internet subscriptions are usually defined as a three-part tariff with a monthly fee Φ, a data cap $\tilde{\theta}$ and an overage fee (i.e. a price for additional data) $\tilde{\phi}$. Therefore a consumer with a consumption equal to θ pays a total amount of $\Phi + \max[0, \theta - \tilde{\theta}]\tilde{\phi}$. In the case of zero-rating, the price of additional data is different for CP_S ($\tilde{\phi}_S$) and CP_W ($\tilde{\phi}_W$), with one of these two prices equal to zero. If content i is zero-rated ($\tilde{\phi}_i = 0$) then the total payment for a consumer choosing to consume from CP_i is equal to Φ, while if he chooses CP_j, it is equal to $\Phi + (\theta - \tilde{\theta})\tilde{\phi}_j$. In our model, consumers have a fixed demand equal to one ($\theta = 1$). With a uniform marginal price (in the case of net neutrality and prioritization), we will use a single price p as a reduced form to represent the subscription paid to the ISP, with $p = \Phi + (1 - \tilde{\theta})\tilde{\phi}$. In the zero-rating case, we will represent the subscription fee by a base price $p = \Phi$ and an additional premium $\phi = (1 - \tilde{\theta})\tilde{\phi}_j$ for consuming the non zero-rated content i.e. content are sold at different prices $p_S \neq p_W$.

2.3. Consumers

The utility of a consumer when he chooses content i depends on (i) the price paid to the ISP for accessing the internet, (ii) his preference for content i over content j, this horizontal differentiation is captured by his position on the Hotelling line, and (iii) the download quality of the content, q_i.

There is a mass one of consumers uniformly located on the Hotelling line. If we denote by x the location of the consumer on the line, by τ the unit transportation cost, by v the gross surplus from being connected to the internet and consuming one unit of content, and by p_i the price paid to the ISP for accessing content i of quality q_i, the utility of consumer x is defined as:

$$U(x) = \begin{cases} v + q_S - \tau|x - a| - p_S & \text{if he chooses } CP_S \\ v + q_W - \tau(1 - x) - p_W & \text{if he chooses } CP_W. \end{cases}$$

[10] Although in the rest of the paper we assume that content providers follow an advertising-based business model, our qualitative results also hold in a subscription-based business model where CPs charge consumers access fees. See Section 8 for details.

[11] The M/M/1 is a queuing model that assumes a single server, arrivals following a Poisson process, and exponential waiting times.

We suppose that consumers have a reservation utility normalized to zero. The following assumption, standard in Hotelling models, guarantees the positivity of equilibrium prices.

Assumption 3. $v \geq \tau$.

2.4. Regulatory regime

We will consider two different regulatory regimes: the net neutrality and the laissez-faire. Under Net Neutrality (NN), both types of discrimination are prohibited. In a neutral internet, we have $q_S = q_W = q$ and $p_S = p_W = p$. In the laissez-faire regime, the ISP can discriminate in price or in quality between different types of content.[12] There are thus two different regimes in a laissez-faire economy: Prioritization (P) and Zero-rating (ZR). Under prioritization, the ISP gives priority to CP_i over CP_j and creates vertical differentiation between the two content $q_i > q_j$. In this regime, the price paid by the consumers to the ISP is the same: $p_S = p_W = p$. Under zero-rating, there is no vertical differentiation between content: $q_S = q_W = q$ but the ISP financially discriminates between the zero-rated content i and the non zero-rated content j: $p_i > p_j$. Under P and ZR, the CPs compete for being privileged by offering a compensation to the ISP.

2.5. Timing of the events and equilibrium concept

The timing of the events is the following:

1. In the laissez-faire regime, the ISP decides to implement P, ZR or nothing (which is equivalent to NN).
2. Under P or ZR, the content providers compete for preferential treatment.
3. The ISP sets the connection prices p_S and p_W.
4. Consumers simultaneously decide which CP to buy from.

In our model, the price charged to consumers and the consumers' choice of content in each regime is independent of the price charged to CPs (either for being prioritized or being zero-rated). We can therefore start with the solution of stages 3 and 4 of the full game and we refer to these as the *consumer pricing subgame*. There are five different consumer pricing subgames: net neutrality, prioritization giving priority to either the strong or the weak firm, and zero-rating the content of either the strong or the weak firm.

Under NN and ZR, the quality of content the consumers face is independent of the action of other consumers. However, under P, the utility of accessing prioritized content decreases in the mass of other consumers subscribing to the priority content. Therefore, when making their decision in stage 4 of the game, consumers are part of a game involving all other consumers. We apply the concept of rational expectations equilibrium to solve this subgame, i.e. we assume that consumers can correctly anticipate the decision of other consumers.[13] From a technical viewpoint, this means that we will find the location of indifferent consumers as fixed points of the demand functions. Note that this assumption is common in the literature (see e.g. Choi and Kim, 2010). As usual, we will solve the full game by backwards induction using the solution concept of subgame-perfect Nash equilibrium.

3. Consumer pricing subgames

In this section we will derive equilibria of the pricing subgames that consist of the ISP's pricing decision followed by the consumers' choice of content.

3.1. Consumer pricing subgame: Net neutrality

Under net neutrality there is a unique price p to access both CPs' content and their quality q are identical. The indifferent consumer between CP_S and CP_W is characterized by:

$$v + q - p - \tau(x^{NN} - a) = v + q - p - \tau(1 - x^{NN}) \quad \Leftrightarrow \quad x^{NN} = \frac{a+1}{2}.$$

Using $q = \kappa$, the ISP maximizes its profit by extracting all the surplus from the indifferent consumer and chooses a price

$$p^{NN} = v + \kappa - \frac{\tau}{2} + \frac{a\tau}{2}.$$

Given Assumption 1, the consumer located at 0 is willing to buy at that price and the market is fully covered. The profits of the ISP and the CPs are equal to $\pi_{ISP}^{NN} = p^{NN}$, $\pi_{CP_S}^{NN} = \rho\frac{a+1}{2}$ and $\pi_{CP_W}^{NN} = \rho\frac{1-a}{2}$.

[12] We consider the possibility to discriminate at the same time in price and in quality in Section 8.

[13] Consumers of the non-prioritized content exert negative externalities on one another and, possibly, positive externalities on consumers of the other content. The solution concept is thus akin to the one used when network effects are present.

3.2. Consumer pricing subgame: Zero-rating

Under ZR, the two content providers have the same download quality $q=\kappa$ but they are financially differentiated. Accessing the zero-rated content costs p while accessing the other costs $p+\phi$. The ISP has the choice between zero-rating the strong or the weak content, and we show below that the latter option is always strictly preferred.

When CP_W is zero-rated, the location of the indifferent consumer x is given by:

$$v+\kappa-p-\phi-\tau(x-a)=v+\kappa-p-\tau(1-x) \quad \Leftrightarrow \quad x=\frac{1+a}{2}-\frac{\phi}{2\tau}.$$

The ISP's maximization problem is the following:

$$\max_{p,\phi} \quad p+x\phi=p+\frac{1+a}{2}\phi-\frac{\phi^2}{2\tau} \qquad \text{s.t.} \qquad p\leq v+\kappa-\tau(1-x).$$

The solution is $p^{ZR}=v+\kappa-\frac{\tau}{2}+\frac{a\tau}{4}$ and $\phi=\frac{a\tau}{2}>0$.

The premium charged by the ISP for accessing the strong content increases with the size of the home market a. The intuition for this result is that the ISP can benefit from reducing the initial asymmetry of CPs (measured by a) by charging a higher price to the strong CP, which in turn means zero-rating the weak CP.

The price charged for zero-rated content is lower than the uniform price under net neutrality, whereas the non-zero-rated content is more expensive: $p^{ZR}<p^{NN}<p^{ZR}+\phi$. These prices lead to the indifferent consumer being located at $x^{ZR}=\frac{1}{2}+\frac{a}{4}$. The average price $\tilde{p}^{ZR}$, i.e. the ISP's profit is:

$$\tilde{p}^{ZR}=\pi^{ZR}_{ISP}=p^{ZR}+x\phi=v+\kappa-\frac{\tau}{2}+\frac{a\tau}{2}+\frac{a^2\tau}{8}=\pi^{NN}+\frac{a^2\tau}{8}.$$

Hence, zero-rating the weak firm is more profitable for the ISP than the neutral regime whenever $a>0$ and the benefits of ZR for the ISP increase with a. It is easy to check that zero-rating the strong firm results in lower profits. Again, Assumption 1 guarantees that the consumer located at 0 is willing to buy so the market is covered. The profits of the CPs are: $\pi^{ZR}_{CP_S}=\rho\frac{2+a}{4}$ and $\pi^{ZR}_{CP_W}=\rho\frac{2-a}{4}$.

3.3. Consumer pricing subgame: Prioritization

With prioritization, the two content providers offer different qualities. Qualities depend on market shares and according to the rational expectations equilibrium concept we assume that in equilibrium consumers correctly anticipate the qualities.

There are three possible equilibrium configurations of the pricing subgame. When content i has priority, we call

1. an *interior equilibrium* when n_i, $n_j>0$ and $(q_i,q_j)=(1,\frac{\kappa-n_i}{n_j})$,
2. a *semi-corner equilibrium* when n_i, $n_j>0$ and $(q_i,q_j)=(\kappa/n_i,0)$, and
3. a *corner equilibrium* when $n_j=0$ and $(q_i,q_j)=(\kappa,0)$.

In a corner equilibrium, there is market tipping in the sense that all consumers choose the prioritized firm. Such an equilibrium arises typically when the transportation cost τ is very low. In a semi-corner equilibrium, some consumers choose the non-prioritized firm even though it has the lowest possible quality. Such an equilibrium arises typically when congestion is very severe, i.e. the capacity κ is small. Finally, there is an interior equilibrium where both market shares and qualities are positive. We will focus our attention on interior equilibria, as it is standard in models of two-sided platforms in general (see Armstrong (2006), and Armstrong and Wright (2007)) and models of net neutrality in particular (see Choi and Kim (2010) and Cheng et al. (2011)). Firstly, we derive conditions of existence of such equilibria, which is equivalent to defining a lower bound on κ. Secondly, we show that when such an equilibrium exists, it is the unique equilibrium. The derivations of the different equilibria are relegated to Appendix B.

If the ISP gives priority to the weak firm's content, in an interior equilibrium the indifferent consumer's location x is given by

$$v+q_S-p-\tau(x-a)=v+1-p-\tau(1-x),$$

where the quality of the stronger CP's content is $q_S=\frac{\kappa-(1-x)}{x}=1-\frac{1-\kappa}{x}$, given that x is indeed an interior solution, i.e. $1-x<\kappa$. Solving this leads to two potential interior solutions for x and the equilibrium consists in selecting the root associated with the highest price. The following lemma describes the equilibrium.

Lemma 1. *An interior solution for the pricing subgame exists under the regime when the weaker CP has priority whenever* $\kappa\geq\kappa^*\equiv\frac{1+\tau-a\tau}{2\tau}$. *It consists of the ISP choosing a price equal to* $p^P=v+1-\tau\left(\frac{3-a}{4}-\sqrt{\frac{(1+a)^2}{16}-\frac{1-\kappa}{2\tau}}\right)$, *with the indifferent consumer located at* $x^P_W=\frac{1+a}{4}+\sqrt{\frac{(1+a)^2}{16}-\frac{1-\kappa}{2\tau}}$.

According to Lemma 1 an interior equilibrium exists if the capacity is sufficiently large. From now on we will assume that this is the case:

Assumption 4. $\kappa \geq \kappa^* \equiv \frac{1+\tau-a\tau}{2\tau}$.

Assumptions 2 and 4 together imply that $\tau(1+a) > 1$ i.e. the products must be sufficiently different from each other to avoid market tipping, a standard assumption in the literature (see e.g. Choi and Kim, 2010, p. 454; and Cheng et al., 2011, p. 65). Therefore, we restrict our attention to situations where congestion is not too severe and content are sufficiently differentiated horizontally.

We can similarly compute the interior equilibrium in which the ISP prioritizes the strong CP and the conditions for the existence of such an equilibrium. We show in the Appendix that whenever an interior equilibrium in the subgame prioritizing the strong CP exists, an interior equilibrium in the subgame prioritizing the weak CP also exists. Moreover, this latter always results in a higher price for the ISP, therefore the former equilibrium can never be a subgame-perfect equilibrium of the full game. The next Proposition establishes an even stronger claim: The interior equilibrium when the weak CP is prioritized dominates all other potential equilibria when P is chosen in the first stage, i.e. it dominates all the potential corner and semi-corner equilibria as well.

Proposition 1. *Under Assumptions 1–4, an interior equilibrium where the ISP gives priority to the weak content exists and it dominates all other potential equilibria of the pricing subgames involving prioritization.*

It follows that under prioritization the equilibrium price and the ISP's profit are:

$$p^P = \pi^P_{ISP} = v + 1 - \tau\left(\frac{3-a}{4} - \sqrt{\frac{(1+a)^2}{16} - \frac{1-\kappa}{2\tau}}\right).$$

At this equilibrium, the indifferent consumer is located at $x^P_W = \frac{1+a}{4} + \sqrt{\frac{(1+a)^2}{16} - \frac{1-\kappa}{2\tau}}$. Assumption 1 guarantees that the consumer located at 0 is willing to buy and the market is covered. The CPs' profits are: $\pi^P_{CP_S} = \rho(1 - x^P_W)$ and $\pi^P_{CP_W} = \rho x^P_W$.

Corollary 1. *It is not in the ISP's interest to exclude any of the content providers from the market.*

In our model, a corner solution corresponds to the ISP creating an extreme quality difference that results in the de facto exclusion of one of the CPs. Indeed, such an exclusion of non-prioritized content is one of the main concerns of net neutrality advocates, both relating to prioritization and to zero-rating (see Marsden, 2010; Van Schewick, 2016). Recall that Proposition 1 ensures the optimality of an interior solution with positive market shares for both CPs and a limited quality differential. Solutions with extreme quality differential (semi-corner equilibria) or market tipping (corner equilibria) are suboptimal. Intuitively, even if the ISP has the ability to implement such a solution, it is not optimal to do so as the price it can extract from consumers would then be lower.

4. Comparisons of regimes without side payments

In this section, we compare the equilibria in the three different regimes without monetary transfers from the CPs to the ISP.

4.1. Market shares

For the consumers, the ISP is a bottleneck access to the content providers. In a neutral internet, the ISP cannot discriminate between content: they are offered by the ISP at the same price and at the same download quality. Competition between content to attract audience is not altered by the ISP. When we depart from neutrality, the ISP has the possibility to influence the market for content by altering the degree of differentiation between content, either using price or quality as instrument.

The price difference under zero-rating is represented by the surcharge ϕ applied to the non-prioritized content. The quality difference between the prioritized and the non-prioritized content is equal to $\Delta q = 1 - q_S$ in the interior equilibrium with priority to the weak. For $\phi > 0$ and $\Delta q > 0$, the privileged content has an extra audience compared to the neutral regime and, as we will discuss in Section 5, the ISP can extract part of this benefit in a previous stage of the game.

In our model, the quality differential Δq, which is a form of vertical differentiation between content, and the price differential ϕ can be easily compared. The indifferent consumers under ZR and P are located at:

$$x^{ZR} = \frac{1+a}{2} - \frac{\phi}{2\tau}, \quad x^P = \frac{1+a}{2} - \frac{\Delta q}{2\tau}.$$

Comparing the indifferent consumers x^{ZR} and x^P which define the market shares of each CP, it is clear that a price differential ϕ has the same impact as an equivalent quality differential Δq. The extra audience gained by the privileged content is larger with quality differentiation than with price differentiation if $\Delta q \geq \phi$ as in this case $x^P \leq x^{ZR}$.

The logic of price discrimination is to charge a higher price for the content that consumers like more. In our model, this preference is captured by the "home market" of the strong content. A higher parameter a (more consumers preferring the strong content), translates into a higher surcharge $\phi = a\tau/2$. This logic applies independently of the network congestion, as the ISP offers the same download quality. Thus the optimal surcharge ϕ does not depend on the capacity κ.

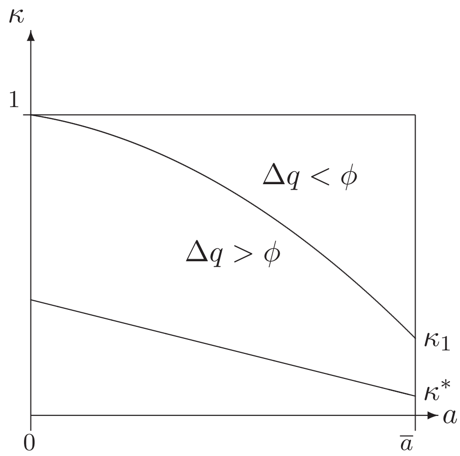

Fig. 2. Price and quality differentiation.

On the contrary, the possibility to differentiate content with respect to download quality depends on the bandwidth capacity κ.[14] With a large bandwidth (κ close to 1), the differentiation between content is limited, while for more severe congestion, differentiation Δq will be larger. In other words, the quality differential Δq is decreasing in κ. Furthermore, we show that:

Lemma 2. *There exists* $\kappa_1 = 1 - \frac{a\tau(2+a)}{8}$ *such that for* $\kappa = \kappa_1$, $\Delta q = \frac{a\tau}{2} = \phi$.

A direct consequence of this result is that $x^P < x^{ZR}$ for $\kappa < \kappa_1$ i.e. prioritization has a greater impact on the market for content than zero-rating when congestion is relatively severe. Conversely, $x^P > x^{ZR}$ holds for $\kappa > \kappa_1$, i.e. for more limited congestion.

The ability to create distortions on the content market depends on the bandwidth and the intrinsic differentiation between CPs. Figure 2 illustrates the results of Lemma 2 in the (κ, a) space showing that for higher values of κ and a, price distortions are more important than quality distortions.

4.2. Prices and profits

We now look at the impact of these distortions on the ISP and the consumers. The prices p^{ZR} and p^{NN} are both linearly increasing in κ with a slope equal to 1. This means that, the benefit of increasing the bandwidth and therefore the quality delivered to consumers can be fully captured by the ISP in a neutral or a zero-rating regime. And, as we have shown earlier, the profit of the ISP is larger under zero-rating than under net neutrality as $\tilde{p}^{ZR} > p^{NN}$. Hence, it always pays for the ISP to zero-rate the weak content.

On the other hand, the price p^P is not linear but strictly concave in κ. With prioritization, quality is not uniform across consumers and content; it is equal to 1 on the fast lane and to q_i on the slow lane. Therefore, should the ISP increase the bandwidth, it will only benefit those who are on the slow lane, i.e. consumers of the non-prioritized content. Furthermore, the benefit of a larger bandwidth for consumers on the slow lane is larger when this extra capacity is shared between few of them i.e. when the non-prioritized content has a lower market share. Since this market share x^P increases in κ, the benefit of increasing capacity is increasing but concave in κ. As the price charged by the ISP reflects the benefit of the marginal consumer, the price is concave too. We will show (Lemma 4) that for exactly the same reason, the consumer surplus is convex in κ.

Comparing the prices $\tilde{p}^{ZR}$ and p^P, we show that the average price is always higher under ZR than under P except for $\kappa = \kappa_1$ where they are equal. This equality is a consequence of Lemma 2; for $\kappa = \kappa_1$ consumers view P and ZR as totally equivalent, hence their willingness to pay is the same. Comparing the prices under NN and P, we find that p^P is larger than p^{NN} if $\kappa \geq \kappa_2 = 1 - \frac{a\tau}{2}$, with $\kappa_2 < \kappa_1$. We summarize our findings in the following lemma:

Lemma 3. *(1) If* $\kappa_2 \geq \kappa^*$, *then* $\tilde{p}^{ZR} \geq p^P \geq p^{NN}$ *for all* $\kappa \in [\kappa_2, 1]$ *and* $\tilde{p}^{ZR} \geq p^{NN} \geq p^P$ *for all* $\kappa \in [\kappa^*, \kappa_2]$. *(2) If* $\kappa_2 \leq \kappa^*$, *then* $\tilde{p}^{ZR} \geq p^P \geq p^{NN}$ *for all* $\kappa \in [\kappa^*, 1]$.

The results of Lemma 3 are represented on Fig. 3. Lemma 3 shows that without side payment from the content providers, the ISP will always implement a zero-rating policy with the weak content being privileged. The intuition for this result is the following. Zero-rating is a tool that the ISP can use to create any level of asymmetry between content it chooses to, therefore the optimal zero-rating offer is the one that creates the optimal level of asymmetry. Prioritization is a more

[14] Assuming that the ISP will not deliberately slow down content when there is capacity available.

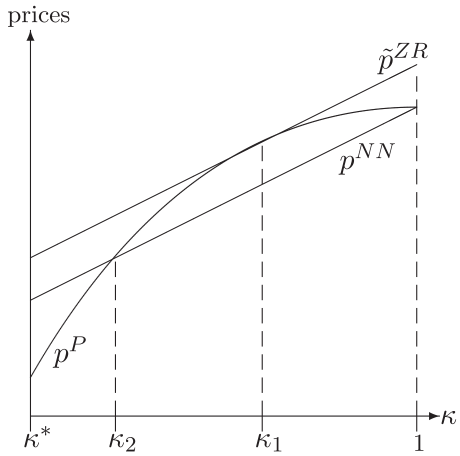

Fig. 3. Prices under NN, ZR and P for $\kappa_2 > \kappa^*$.

constrained tool for the ISP as the quality differential is endogenously determined by the consumers' choices. Indeed, as Lemmata 2 and 3 show, for $\kappa < \kappa_1$ the asymmetry created by prioritization is higher than optimal, whereas for $\kappa > \kappa_1$ it is lower than optimal.

5. Potential transfers from the CPs to the ISP

In the previous section we search for the preferred policy for the ISP, ignoring that CPs can compensate the ISP for having a privileged treatment. In this section, we solve the full game and investigate how the ISP's ability to extract money from the CPs alters market outcomes.

Importantly, we assume that in the beginning of the game the ISP credibly commits to offering only one type of preferential treatment: either paid prioritization or zero-rating (or nothing), and this becomes common knowledge among all players. Therefore, after committing to paid prioritization, it cannot threaten a CP to approach its competitor with a zero-rating offer and vice versa. In other words, we restrict the ISP's possible threats in the bargaining stage, simplifying the analysis considerably. Given the regime chosen at stage 1, CPs are competing for preferential treatment by simultaneously submit offers to the ISP who then either accepts one of the offers or rejects both. In the latter case the neutral regime applies. An offer consists of a transfer from the CP to the ISP and does not include a commitment to the price charged to consumers. Therefore, the ISP subsequently plays one of the price subgames described in Section 3.

5.1. Zero-rating offers

In the previous analysis we established that in the absence of transfers from CPs to the ISP, the optimal zero-rating policy of the ISP is zero-rating the weak CP. It also clearly follows that the ISP will never benefit from zero-rating the strong CP. In particular, in case the bargaining ends with zero-rating the strong content (corresponding to $\phi < 0$), in the subsequent price subgame the ISP will choose a value for the discount ϕ as close as possible to zero. This means that zero-rating the strong content is equivalent to the NN situation.

Therefore, the credible threat point of the ISP in case its offer is refused by the weak CP is the NN situation. Given that the aggregate demand is fixed, zero-rating only displaces demand from one CP to the other. Hence, the maximal willingness to pay for the preferential treatment is equal to the value of the additional traffic which is equal to $\rho(x^{NN} - x^{ZR}) > 0$ for both CPs. As $p^{NN} < \tilde{p}^{ZR}$, the ISP loses revenue from the consumers if it accepts the offer of the strong firm, thus the equilibrium outcome will be the ISP zero-rating the weak firm.

Next, we determine the optimal bid. There are two cases depending on the size of the CPs' advertising revenue, ρ. First, if it is sufficiently large that the maximal bid of the strong CP compensates the ISP for the loss of revenue, i.e. if

$$\rho(x^{NN} - x^{ZR}) + p^{NN} > \tilde{p}^{ZR} \quad \Leftrightarrow \quad \rho > \frac{\tilde{p}^{ZR} - p^{NN}}{x^{NN} - x^{ZR}} \equiv \rho^{ZR} = \frac{a\tau}{2}$$

then the weaker firm must bid a positive amount to get the zero-rating agreement. In particular, in equilibrium the weak CP's bid equals

$$\rho(x^{NN} - x^{ZR}) + p^{NN} - \tilde{p}^{ZR} > 0$$

which is positive but smaller than its maximal willingness to bid as $\rho > \rho^{ZR}$.

Second, if advertising revenues are relatively small, $\rho \leq \rho^{ZR}$, then the inequality above is not satisfied, thus the strong CP's maximal bid is insufficient to compensate for the ISP's loss of profits, therefore the weak firm gets zero-rated for free, and its optimal bid is 0.

Generally, the weak CP bids $\max\{0;\ \rho(x^{NN} - x^{ZR}) + p^{NN} - \tilde{p}^{ZR}\}$ and wins the zero-rating contract. The ISP's overall profit is given by $\max\{\tilde{p}^{ZR};\ \rho(x^{NN} - x^{ZR}) + p^{NN}\}$.

5.2. Paid prioritization offers

From the previous section we know that under $\kappa \geq \kappa_2$ prioritizing the weak CP is the best prioritizing option for the ISP as it extracts more revenue from the consumers' side. If instead, the ISP prioritizes the strong CP, it will charge a lower price $p_S^P < p^P$ to consumers but the strong CP will gain additional traffic $x_S^P - x_W^P > 0$.[15] However, the traffic gains by the strong CP are the traffic losses of the weak CP. Therefore, the maximal bid of both CPs is given by $\rho(x_S^P - x_W^P)$ and, given that consumers pay more when CP_W is prioritized, the ISP will accept the offer from CP_W.

Following the logic laid out above in the zero-rating case, we first determine the strong firm's maximal bid. Its maximal bid equals $\rho(x_S^P - x_W^P)$. The weak CP must then bid

$$\max\{0;\ \rho(x_S^P - x_W^P) + p_S^P - p^P\}$$

to win the contract, and the ISP's profit is given by $\max\{p^P;\ \rho(x_S^P - x_W^P) + p_S^P\}$.

Finally, $p^P < p^{NN}$ for $\kappa^* \leq \kappa \leq \kappa_2$, thus the ISP prefers neutrality in the absence of a sufficiently large payment from one of the CPs, i.e. it will refuse the offers made by the CPs. Consequently, in order to be prioritized, the weak CP must overturn both neutrality and priority to the strong CP, and it will be able to do so when traffic has a sufficiently high value.

5.3. ISP's choice between paid prioritization and zero-rating

Next we investigate the ISP's decision in a regulatory environment where the ISP is free to choose either prioritization or zero-rating.

The level of CPs' advertising revenues is a key factor in the ISP's choice. From the previous results it clearly follows that for low advertising revenues the ISP prefers zero-rating as $\tilde{p}^{ZR} > p^P$. Intuitively, with low advertising revenues, the CPs cannot afford to compensate the ISP for switching to the other regime, as it would result in relatively large losses on the consumer side.

Conversely, with high advertising revenues the large payments from CPs to the ISP can dwarf the eventual loss of ISP profit from the consumer side when it changes to paid priority programs. Under certain conditions, prioritization creates a larger divide in market shares than zero-rating programs, making them more attractive for the ISP when advertising revenues are high. The next Proposition formalizes these findings.

Proposition 2. *There exist $\kappa' \in (\kappa_1, 1)$ and $\overline{\rho}(\kappa) > 0$ such that prioritization is the equilibrium choice of the ISP for $\kappa < \kappa'$ and $\rho \geq \overline{\rho}(\kappa)$. Otherwise the ISP chooses zero-rating in equilibrium.*

Intuitively, to overturn the zero-rating regime, prioritization must create a larger distortion in market shares than zero-rating. For that, the quality differential between the prioritized and the non-prioritized content should be large enough, which corresponds to a relatively low value of capacity. In addition, traffic should be sufficiently valuable for CPs, as captured by the level of advertising revenues.[16] As demonstrated in Lemmata 2 and 3, without side-payments, i.e. with $\rho = 0$ the ISP benefits from *reducing the asymmetry* of content types as it allows it to charge a higher price on the consumer side. In turn, Proposition 2 shows that for high levels of advertising revenues the ISP prefers trading off lower revenue from the consumer side for higher revenue from the CP side by *increasing the asymmetry* of content types.

6. Welfare analysis

In this section, we focus on the consequences of the ISP's choice on consumer surplus. Total welfare, defined as the sum of consumer surplus, the profit of the ISP and the CPs, is arguably of lesser interest in our framework.[17]

---

[15] In Lemma 5 in the Appendix, we show that when priority to the strong is given, the highest profit is achieved at the interior equilibrium whenever this equilibrium exists, that is when $\kappa \geq \frac{1+\tau+\alpha\tau}{2\tau}$. Notice that this is stronger than Assumption 4, but they are qualitatively similar: they both require a relatively high level of capacity and horizontal differentiation. For the subsequent analysis, we will assume that this condition holds but as we discuss below, our results do not qualitatively depend on this parameter restriction.

[16] This logic is independent of the type of equilibrium under prioritization. In particular, even if an interior solution with priority to the strong CP does not exist, prioritization will be implemented provided that it leads to a sufficiently strong distortion in market shares combined with high value for traffic.

[17] As prices are pure transfers, and average quality is the same in our model with unit consumption and covered market, welfare is highest in the regime where the transportation costs are lowest, that is under net neutrality. This implies that discrimination of content by the ISP - either in price or in quality- is never welfare-improving. This result does not come as a surprise, as in our model, discrimination is only a tool for the ISP to alter competition on the market for content and does not have added value. Comparing P and ZR, the solution leading to higher welfare is the one where transportation costs are lower, i.e. where the market share of CP_W is smaller (see Lemma 2 for details).

6.1. Characterization of consumer surplus

Before we formally compare consumer surplus under the three regimes, it is important to recall three features of the model. First, the average quality is the same in all the configurations we considered (and equal to κ), and in particular, there is no improvement in the average quality under prioritization. Second, under NN and ZR, prices are linear in κ, so the ISP fully captures any quality increase. Last, transportation costs are minimized if there is no discrimination between content types, i.e. they are the lowest under net neutrality. These observations lead to the immediate conclusion that the consumer surplus is higher under NN than under ZR. Formally, the consumer surplus under both NN and ZR is independent of κ and equal to

$$CS^{NN} = v + \kappa - p^{NN} - \int_0^a \tau(a-t)\mathrm{d}t - \int_a^{x^{NN}} \tau(t-a)\mathrm{d}t - \int_{x^{NN}}^1 \tau(1-t)\mathrm{d}t = \tau\left(\frac{1}{4} - \frac{3}{4}a^2\right), \quad \text{and}$$

$$CS^{ZR} = v + \kappa - p^{ZR} - \int_0^a \phi + \tau(a-t)\mathrm{d}t - \int_a^{x^{ZR}} \phi + \tau(t-a)\mathrm{d}t - \int_{x^{ZR}}^1 \tau(1-t)\mathrm{d}t = CS^{NN} - \tau\frac{3}{16}a^2.$$

Consumer surplus under P simplifies to

$$CS^P = v + \kappa - p^P - \int_0^a \tau(a-t)\mathrm{d}t - \int_a^{x^P} \tau(t-a)\mathrm{d}t - \int_{x^P}^1 \tau(1-t)\mathrm{d}t.$$

Contrary to the other cases, the consumer surplus is not constant in κ, as x^P is concave in κ. The following Lemma summarizes the comparison of CS^P to the other two regimes.

Lemma 4.

(i) *The consumer surplus under P is convex in κ with $CS^P = CS^{NN}$ for $\kappa = 1$ and $\frac{\partial CS^P}{\partial \kappa} = 0$ for $\kappa = \kappa_2$.*

(ii) *There exists a cutoff value of capacity $\kappa_3 = 1 - a\tau + a^2\tau$ such that if $\kappa_3 \le \kappa^*$, then net neutrality is the preferred regime of consumers for all capacity levels, i.e. for all $\kappa \in [\kappa^*, 1]$. If $\kappa_3 > \kappa^*$, then $CS^P \ge CS^{NN}$ for $\kappa \in [\kappa^*, \kappa_3]$, i.e. consumers prefer prioritization for low capacity levels; whereas $CS^P \le CS^{NN}$ for $\kappa \in [\kappa_3, 1]$, i.e. they prefer net neutrality for higher capacity levels.*

(iii) $CS^{ZR} \ge CS^P$ if and only if $\kappa_4 = \frac{8-6a\tau+3a^2\tau}{8} \le \kappa \le \kappa_1 = \frac{8-2a\tau-a^2\tau}{8}$

(iv) $\kappa_3 < \kappa_4 < \kappa_2 < \kappa_1$.

The convexity of consumer surplus in part (i) is the direct consequence of the concavity of x^P, as established under Proposition 1. One can find κ_3 defined in part (ii) as the smaller root of the equation $CS^P = CS^{NN}$. Clearly, the larger root must equal 1 as for $\kappa = 1$ prioritization coincides with net neutrality. Intuitively, comparing the surplus under NN and P, it is clear that for all $\kappa_2 \le \kappa < 1$, $CS^{NN} > CS^P$ as the price and the transportation cost are both higher under P. However, for lower values of κ, prioritization leads to a lower price but still a higher transportation cost. Part (ii) captures this tradeoff. The values of intersections in part (iii) are again a consequence of CS^P's convexity. Finally, the order of variables in part (iv) can be checked in a straightforward manner and helps us to plot the relation of consumer surpluses. Figure 4 summarizes these results.

6.2. Consumer surplus in a laissez-faire regime

Proposition 2 describes the optimal policy choice of the ISP in a laissez-faire regime, i.e. in a world where transfers from CPs to the ISP are allowed. The consequences of this choice for consumers can be easily inferred from Fig. 4.

We observe that, when congestion is limited, zero-rating is always harmful to consumers independently of ρ. Note that for large values of advertising revenues, ρ, zero-rating is implemented precisely when congestion is limited. Hence, for $\rho \ge \overline{\rho}$, i.e. when zero-rating is chosen by the ISP, it is the worst regime for consumers (as $\kappa' \ge \kappa_1$).

Next, when prioritization is implemented, it benefits consumers if and only if congestion is severe ($\kappa \le \kappa_3$). In those cases the price charged to consumers is low, and the ISP creates a large distortion on the content market. Hence, it manages to extract a large payment from the prioritized CP. The lower revenues on the consumer side are more than compensated by higher transfers on the CP side. In that sense, the CP 'subsidizes' the consumers who thus benefit from this policy. We summarize this discussion in the following corollary.

Corollary 2. *Regarding the choice between zero-rating, prioritization and net neutrality, for low congestion levels ($\kappa \ge \kappa'$), the interests of the ISP and the consumers are opposed; for severe congestion ($\kappa \le \kappa_3$), the interests of the ISP and the consumers are congruent if and only if advertising revenues are large, i.e. for $\rho \ge \overline{\rho}$.*

7. Larger asymmetry in content types

In this Section we test the assumptions of limited asymmetry we have made in the baseline model in three ways. In the first two models, we relax the fixed demand and fully covered market assumptions, allowing the strong content provider to

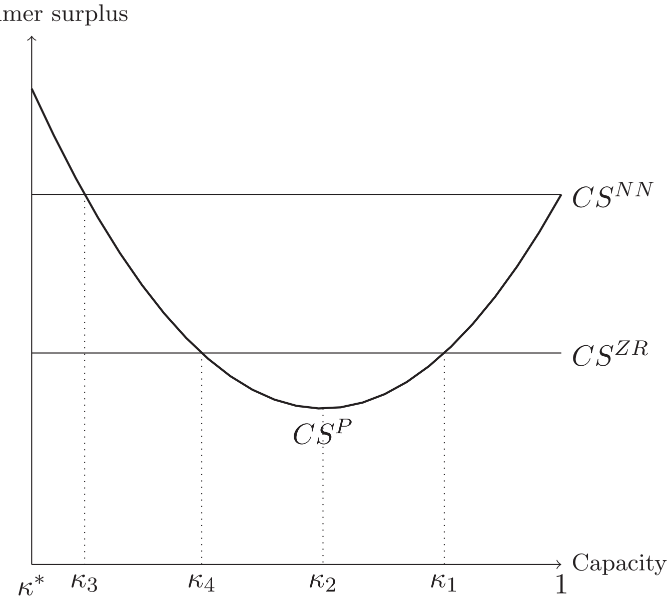

Fig. 4. Consumer surplus under NN, ZR and P for $\kappa_3 \geq \kappa^*$.

be significantly more appealing to consumers than its weak rival. In the third model, we relax the homogenous advertising revenues assumption, allowing the strong CP to monetize traffic on its website better than its rival, thus making it more attractive for the ISP.

7.1. Zero rating and variable demand

The baseline model is a fixed demand model where the market is fully covered and each consumer has a fixed demand for content normalized to one. This is one of the assumptions that implicitly provide a limit to the asymmetry between content providers. In this context, zero-rating and prioritization displace consumers from one content provider to the other but the aggregate demand remains constant. As a consequence, the ISP prefers to make the weak content more attractive in order to extract a higher subscription price from consumers. Whether it prefers prioritization to zero-rating depends on the distortions it creates on the content market and the value of traffic for the CPs (Proposition 2). When the aggregate demand is not constant, either because consumers intensify their consumption of the zero-rated content or because the market is not covered, the ISP may prefer to zero-rate the strong content instead of the weak because it has a higher ability to generate additional traffic. We propose two models with variable demand.[18] In the first, the consumption of content increases when the content is zero-rated. In the second, the market is no longer covered and zero-rating might be used to increase market coverage. Finally, even if we do not formally develop the model, we argue that a similar mechanism is in play with prioritized content.

7.1.1. Variable content consumption

In this section, we suppose that a marginal price equal to zero increases the individual consumption of the zero-rated content. This has three consequences. First, zero-rating generates additional content consumption, directly increasing the consumer surplus.

Second, the increase in demand generates additional traffic which, in a congested network, is costly for the ISP. For a fixed capacity κ, the additional traffic decreases the delivered quality q which in turn decreases the subscription price p. Or, if the capacity is endogenous, the ISP must invest in its infrastructure to cope with the additional traffic. In both cases, it generates an additional cost for the ISP.

[18] A third extension would be to consider multi-homing consumers. Consider a version of our model where an endogenous fraction of the consumers multi-home. By definition, prioritization and zero-rating would make the privileged content relatively more attractive and the non-privileged one less attractive. Consequently, for multi-homers the single-homing option at the privileged content would become relatively more beneficial. We can therefore expect that some multi-homers would switch to single-homing at the privileged content. As a consequence, we should observe less multi-homing in a non-neutral network compared to a neutral one. This process would have a negative impact on the non-privileged CP's revenue, thus CPs would be ready to pay more for being prioritized or zero-rated. Overall, multi-homing would make prioritization and zero-rating more powerful tools for the ISP to manage asymmetry on the content market. Although its differential effects in the regime choice of ISP are a priori unclear, we do not see any obvious reasons why it would change our main results qualitatively.

Fig. 5. The extended Hotelling line.

Third, in our initial formulation when CPs are bidding to be zero-rated, both firms bid the same amount as the traffic gained by one firm is lost by the other. With variable demand, this is no longer the case as zero-rating generates a traffic expansion both at the intensive and at the extensive margin. Therefore, the strong CP can potentially gain more from being zero-rated than the weak one and the ISP may prefer to zero-rate the strong content.

To formalize this last claim, let us consider the following modified model. The ISP still charges a connection price p to consumers and a variable fee per unit of content consumed, ϕ_S for the strong content, ϕ_W for the weak content. These variable fees can take two values, zero or ϕ, which we consider to be given. Each consumer of content i will have a demand equal to 1 if $\phi_i = \phi$ and equal to $1+\theta$ if content i is zero-rated i.e. if $\phi_i = 0$. The individual consumption of the zero-rated content increases by θ, giving an extra surplus Θ to the consumer while the consumption of the non-zero rated content remains constant and equal to one.

The total traffic and therefore the quality q depends on which content is zero-rated. When the strong content is zero-rated, the indifferent consumer is characterized by

$$\tilde{x}^{ZR=S} = \frac{1+a}{2} + \frac{\Theta}{2\tau} + \frac{\phi}{2\tau}$$

The total traffic is equal to $X^{ZR=S} = 1 + \theta\tilde{x}^{ZR=S}$ and the associated quality is $q^{ZR=S} = \frac{\kappa}{X^{ZR=S}}$. Similarly, when the weak content is zero rated, we have

$$\tilde{x}^{ZR=W} = \frac{1+a}{2} - \frac{\Theta}{2\tau} - \frac{\phi}{2\tau}$$

The total traffic is equal to $X^{ZR=W} = 1 + \theta(1-\tilde{x}^{ZR=W})$ and the associated quality is $q^{ZR=W} = \frac{\kappa}{X^{ZR=W}}$.

Zero-rating the strong content generates more traffic ($X^{ZR=S} > X^{ZR=W}$) and, consequently, the associated quality is lower ($q^{ZR=S} < q^{ZR=W}$). The optimal price p is such that the indifferent consumer has a zero utility. Comparing the associated profit, it is straightforward to show that $\pi^{ZR=S} - \pi^{ZR=W} = -a\phi - (q^{ZR=W} - q^{ZR=S}) < 0$, i.e. the ISP has a larger revenue from consumers if it zero-rates the weak content, similarly to the baseline model.

Considering the competition between CPs to be zero-rated, however, CPs are no longer ready to pay the same amount as the total demand is no longer fixed. The bid of the strong CP will be at most equal to the additional revenue from traffic it can expect if it is zero-rated instead of the weak CP. The highest possible bid for the strong CP is thus equal to

$$b^S = \rho(\tilde{x}^{ZR=S}(1+\theta) - \tilde{x}^{ZR=W}) = \rho(\tilde{x}^{ZR=S} - \tilde{x}^{ZR=W}) + \rho\theta\tilde{x}^{ZR=S}$$

And similarly for the weak CP:

$$b^W = \rho[(1-\tilde{x}^{ZR=W})(1+\theta) - (1-\tilde{x}^{ZR=S})] = \rho(\tilde{x}^{ZR=S} - \tilde{x}^{ZR=W}) + \rho\theta(1-\tilde{x}^{ZR=W})$$

Combining the two leads directly to $b^S - b^W = \rho a\theta > 0$. Hence, we can establish that if ρ is large enough, the ISP will choose to zero-rate the strong CP because then the benefit of additional traffic paid by the CP more than compensates the lower revenue paid by the consumers.

Proposition 3. *There exists* $\tilde{\rho} = \frac{\phi}{\theta} + \frac{q^{ZR=W} - q^{ZR=S}}{a\theta}$ *such that, if* $\rho \geq \tilde{\rho}$, *the ISP zero-rates the strong content and if* $\rho \leq \tilde{\rho}$ *then it zero-rates the weak content.*

The Proposition shows that there always exists a range of parameters for which zero-rating the weak CP remains the ISP's optimal choice. This result obtained in the baseline model is then robust.

7.1.2. Non-covered market

Next, we analyze the case in which the market is not fully covered but we suppose that each consumer has a unit demand for content. For that, we consider an extended Hotelling line with a mass b of additional consumers located at the left of the original interval. To avoid dealing with negative numbers for the consumers' location, let us suppose that the Hotelling line has a length of $1+b$, CP_W is located at the right extreme ($x=2$) and CP_S is located at $x = 1+a$, and the most extreme potential consumer sits at $1-b$ (see Fig. 5). Clearly, the asymmetry between CPs is increasing in the parameter b, the mass of additional consumers closer to the strong CP.

In this case, there are two indifferent consumers: the consumer $\tilde{x}$ who is indifferent between CP_S and CP_W, with $\tilde{x} \in [1+a, 2]$ and the consumer $\underline{x}$ who is indifferent between CP_S and not subscribing to the ISP, with $1-b \leq \underline{x} \leq 1+a$. The market shares of CP_S and CP_W are respectively $(\tilde{x} - \underline{x})$ and $(2-\tilde{x})$.

The ISP chooses the base subscription price p and decides which content is zero-rated. As in the previous extension, we denote the additional price for accessing CP_S by ϕ. When $\phi > 0$, the weak content is zero-rated, when $\phi < 0$, it is the strong content that is zero-rated. The maximization problem of the ISP is the following:

$$\max_{p,\phi} (2-\underline{x})p + (\tilde{x} - \underline{x})\phi \quad \text{subject to} \quad U(\tilde{x}), U(\underline{x}) \geq 0; \; \underline{x} \geq 1-b$$

Proposition 4. *There exists b_0, b_1 and b_2 with $0 < b_0 < b_1 < b_2$ such that: (i) When $b \le b_2$, $\underline{x} = 1 - b$ and when $b \ge b_2$, then $\underline{x} = 1 - b_2$. (ii) The optimal surcharge ϕ^* increases in b for $b \in [0, b_0]$ and decreases for $b \in [b_0, b_2]$. (iii) The ISP optimally zero-rates CP_W ($\phi^* \ge 0$) when $b \le b_1$ and CP_S ($\phi^* \le 0$) when $b \ge b_1$.*

Proposition 4 states that, with a variable demand, the market will be covered up to $1 - b_2$. The price surcharge ϕ is used to extract surplus from consumers of the strong content but also to attract consumers on the left of the line. For this reason, the optimal surcharge ϕ^* is concave in b, increasing for low values of b to collect more surplus and decreasing for middle values of b to extend the market size. At a point, it even becomes negative and the ISP optimally starts zero-rating the strong CP. Accordingly, part (iii) shows that the ISP realizes the highest profit, including the potential bids received from the CPs, by zero-rating the weak CP for $b \le b_1$ and the strong CP otherwise. In this variable demand model, the ISP will zero rate the strong content only when its capacity to increase demand is very significant, i.e. when the asymmetry between content types is very large. Otherwise, it zero rates the weak CP as in the baseline case.

To conclude, a similar argument can be made for prioritization of content. In a fixed demand model, the ISP always prefers to prioritize the weak content in order to extract a higher price from consumers. When demand is variable, by the same logic as above, there exists a range of additional demand where prioritizing the weak content remains profitable. When the potential gain of market share is sufficiently large, however, the ISP may give priority to the strong CP as it is a tool to enlarge its demand. Furthermore, with variable demand and fixed capacity, prioritization is likely to create a large quality gap between the two CPs, and, as we have shown, it is exactly in those circumstances that the ISP can extract the largest amount from the CPs. Therefore, we conjecture that the results presented in Proposition 2 can be extended to the non-covered market case.

7.2. Heterogeneous advertising revenues

In our model, we call the strong CP the one that has a larger market share without discrimination by the ISP. Another potential difference between the CPs is their capacity to monetize traffic (as in Bourreau et al., 2015; Jullien and Sand-Zantman, 2018; Somogyi, 2018). A plausible reason for having a higher ability to monetize the traffic for the strong CP is that a higher consumption level implies more data, which can entail more revenue through a better targeting of the ads or data trade and partnerships. In this extension, we investigate that case by considering $\rho_S > \rho_W$ and we show that our results are robust to this alternative specification.

In Section 5 we show that the choice of regime depends both on the valuation of traffic by the CPs and the payment the ISP can collect from the consumer side. In equilibrium, the weak CP pays less than the strong CP would be willing to pay and still secures preferential treatment because the ISP can collect more revenue from consumers when giving preferential treatment to the weak CP.

With heterogeneous advertising revenues, when traffic is more valuable for the strong CP, its bid increases relative to the bid of the weak CP. However, to overturn the regime choice of the ISP, this bid increase must be sufficiently large to compensate for the losses on the consumer side. Therefore ρ_S must be substantially larger than ρ_W to overturn the market equilibrium. In this sense, our results are robust: a limited difference between advertising revenues is not sufficient for the strong CP to gain preferential treatment.[19]

8. Extensions and robustness checks

In this Section we briefly summarize extensions and robustness checks of our baseline model.

Subscription-based business model. In the baseline mdoel we assume that content providers follow an advertising-based business mode. However, our qualitative results also hold in a subscription-based business model where CPs charge access fees to consumers. To see this, we introduce an additional first stage in the model where CPs simultaneously choose the level of subscription fees.[20] We show that despite the strong CP's ability to directly benefit from its advantageous position (by charging a higher fee to consumers than its rival), the ISP still has room to create additional profitable differentiation by zero-rating the weak CP. Therefore, the logic of the advertising business model described above still applies to the subscription-based business model. See Appendix A.1 for details.

Simultaneous prioritization and zero-rating. In the baseline model, the ISP either prioritizes content, or zero-rates it, or does neither. In this extension we study the optimal policy of the ISP when it has the opportunity to engage in both practices at the same time. In other words, the ISP can discriminate both in price *and* in quality.

The ISP can zero-rate the non-prioritized or the prioritized content and there are real-life examples of both. In practice, ISPs have created such situations by throttling specific content. An example of zero-rating of non-priority content is T-Mobile's Binge On program where the zero-rated video content is not available in HD (the resolution is limited to 480p). At the same time, in Europe there have been examples for zero-rated priority content. Some ISPs offered zero-rating contracts

[19] In Reggiani and Valletti (2016), the probability of clicking on the content's ad-banner is larger for the prioritized content. In this case, we would have $\rho_S < \rho_W$ in the prioritization regime which would increase the potential payment from the weak content. This would clearly reinforce our results.

[20] In addition to being very tractable, this timing is the one consistent with the covered market assumption, as demonstrated by Jeitschko et al. (2018).

where the traffic was slowed down or eventually interrupted for all content but the zero-rated one when the monthly data cap was reached.[21] Thus, the price and quantity discrimination are not necessarily going in the same direction.

Using our model, we first show that when combined with zero-rating, the ISP prefers to give priority to the weak content in an interior equilibrium, similarly to the baseline case. For relatively high levels of capacity, the quality differential between the two content is limited. In this case, the ISP zero-rates the weak content and price discrimination reinforces quality discrimination. On the contrary, for relatively low levels of capacity, quality differentiation is important. In this case, the ISP prefers to zero-rate the strong content and price discrimination mitigates quality discrimination.

However, we also find that the profit with prioritization and zero-rating is lower than the profit with zero-rating alone as long as there is some congestion. Thus, without side-payments, the ISP will use just one out of the two available instruments, the price, to discriminate between content. This is reminiscent of the results of Economides (1989), and Neven and Thisse (1989) who show that firms choose maximal differentiation in one dimension (the price in our model) and minimal differentiation in the other (the quality in our model).

With side-payments from CPs, simultaneous prioritization and zero-rating can be advantageous for the ISP whenever advertising revenues are sufficiently high. We provide a sufficient condition for this. See Appendix A.2 for details.

Investment in capacity. In the baseline model, we have assumed that the capacity of the ISP is exogenous. In this section we relax this assumption and let the ISP invest in costly capacity building at an initial stage of the game. Investment incentives of different regulatory regimes on the internet are an important aspect of both the policy debate and the academic literature (see e.g. Choi and Kim, 2010; Economides and Hermalin, 2012; Bourreau et al., 2015). On the one hand, opponents of net neutrality claim that the additional revenue from the CP side is necessary for the ISP to finance capacity building. On the other hand, advocates of net neutrality fear that allowing preferential treatment creates perverse incentives for the ISP to keep capacity levels low in order to continue extracting benefits from the CP side.

Our analysis of investment incentives under prioritization and our results are similar to the analysis of Choi and Kim (2010). There are two effects in play. First, prices under prioritization are strictly concave in capacity κ i.e. additional capacity only improve the delivery speed on the slow lane. Second, congestion reinforces the distortion on the content market created by prioritization. The lower the capacity levels, the higher the distortion is, and consequently the higher the level of side payments the ISP can extract from the prioritized CP. Therefore, under prioritization the ISP may have an incentive to sustain an inefficiently high level of congestion in order to extract more payments from the CP side. This happens when advertising revenues are sufficiently large, which is precisely the condition for the ISP implementing prioritization (see Proposition 2). Thus prioritization may create a trap for the ISP in the sense that a medium initial capacity level leads to reduced subsequent investment, especially for high advertising revenues, reflecting the fear of net neutrality advocates. Thus, under prioritization, the capacity level chosen by the ISP is generically socially suboptimal.

On the contrary, in the zero-rating regime, we find that the ISP's investment incentives are completely aligned with the social optimum as the ISP's profit are simple linear functions of the capacity level. Additional capacity speeds up uniformly the content delivery and the ISP can charge more for access. Intuitively, zero-rating, a form of price discrimination, does not require congestion to be profitable. See Appendix A.3 for details.

Competing ISPs. In the baseline model we consider a monopolistic ISP. However, internet service providers potentially compete on both sides of the market. On the user side, they compete to attract consumers. On the content side, they compete to privilege attractive content. Let us consider competing ISPs, each of them having an installed base of users i.e. we focus first on competition on the content side of the market. As their clients do not multihome, the ISPs offer privileged access to their client base to the CPs. Each ISP is thus the unique provider of additional traffic from its clients to the competing content providers. Therefore, the CPs compete for privileged access at each ISP. In our baseline model, we considered competition between CPs for being prioritized or zero-rated at a single ISP. But this model can be easily transposed to the case of competing ISPs as each of them offers monopolistic access to its clients. Therefore, we believe that our specification of competition on the CP side easily extends to the case of competing ISPs.

Competition on the user side is more complicated. Our model shows that discrimination of content generates additional income for the ISP when the traffic is sufficiently valuable for the content providers. Therefore, one could expect that the additional profits generated on the content side intensify competition on the user side. The first and most obvious effect of competition between ISPs is a likely price decrease for users, its size depending on the elasticity of demand. However, if the price elasticities of demand are limited - as it is the case in our baseline model– competition on the user side is limited, and our model and its results can be used to describe the case of competing ISPs. This would be the case, for instance in models of horizontally differentiated ISPs as in Bourreau et al. (2015) or Choi et al. (2015) when consumers have a strong preference for one ISP over the other.

However, ISPs do not only compete in prices to attract users. Offering privileged access to content is part of the competitive process. On the one hand, the ISPs want to offer privileged access to the most valuable content, which is the one that can attract the largest revenue from the CP or the user side. On the other hand, ISPs want to differentiate themselves from their competitors, a reason for not privileging the same content. It is, therefore, a priori not clear whether competing ISPs will privilege the same content. Broos and Gautier (2017) consider a similar framework where competing ISPs can include

[21] These contracts with zero-rating and throttling were considered to be a violation of the European net neutrality regulations and later forbidden (see Krämer and Peitz, 2018 on these points).

or not a competing application in their bundle of products. They show that if an ISP excludes the app, the other has incentives to offer it, i.e. app exclusion at both ISPs is not an equilibrium. But both ISPs offering the app could be an equilibrium. Similarly, we conjecture that, in our context, an ISP always has incentives to propose a privileged offer if the other does not. However, our model cannot predict whether all operators will propose zero-rating offers (as in Portugal) or only some of them (as in Germany or the UK). Neither can we conclude that offers will be different, as in the US where mobile operators zero-rate different video content, or similar like in Portugal where the three leading operators zero-rate the same bundle of applications.[22]

9. Discussion

9.1. Policy implications

In this section, we highlight two policy interventions for which considering the three regulatory regimes within the same model is crucial to drawing the right conclusion about the implications for consumer surplus. To put these potential interventions in context, we start by describing current regulatory approaches in different jurisdictions.

Current regulatory approaches to zero-rating and prioritization

In the US, the 2015 Open Internet Order explicitly mandates a case-by-case treatment of zero-rating, i.e. zero-rating was tolerated, but in contrast, paid prioritization was prohibited by a bright line rule. The FCC, the US telecommunications regulatory body, investigated the four zero-rating offers in place in 2016 and expressed concern over two of them (Verizon's FreeBee Data and AT&T's Sponsored Data programs) likely harming consumers.[23] These two programs have in common that the ISP zero-rated exclusively its affiliated video service subsidiary, while the other programs such as Binge On by T-Mobile were non-exclusive and multiple content providers were eligible for zero rating. However, this investigation did not have any legal effect as the FCC's composition changed under the Trump administration, and all further investigation was stopped. Moreover, the new FCC voted a repeal of the 2015 rules and abolishing net neutrality protections in December 2017. The repeal has resulted in both business practices being currently permitted in the US.

In 2015, the EU decided to adopt rules similar to the US regulation that were in place between 2015 and 2017. 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, while explicitly banning paid prioritization. The Guidelines explicitly prohibit zero-rating of content in case the data cap is reached and all other content is throttled or blocked. In practice, according to this principle, some specific zero-rating offerings were banned by national regulatory authorities (see e.g. the case of Telenor in Hungary, Telia in Sweden and T-Mobile in Germany). In contrast, other offerings were found legal (e.g. Proximus in Belgium).[24] While Canada also decided not to ban zero-rating ex-ante, in practice CRTC, its regulatory authority took a very strict approach toward it, de facto prohibiting both zero-rating and paid prioritization.[25] For a detailed summary of the history of zero-rating regulation up to the end of 2016, see Yoo (2016).

Asymmetric regulatory approach banning prioritization

As described above, some regulators have an asymmetric approach to prioritization and zero-rating, banning the former, allowing the latter under some conditions. Previous work on net neutrality regulation compared either prioritization programs or zero-rating programs to net neutrality. However, our model points out that this can be misleading, as the relevant choice of the ISP is not between prioritization and neutrality but between prioritization and zero-rating.

When congestion is limited, and the asymmetry between CPs is not very large, our model shows that the ISP prefers zero-rating to prioritization because price discrimination is more effective to extract surplus from consumers and content providers. For high values of κ, prioritization has a limited impact on consumers and on the market for content. A ban on prioritization has then little effect as the ISP prefers zero-rating anyway.

When congestion is more severe, prioritization creates more distortion on the market for content and on the incentives to develop the infrastructure. This however must be traded-off with a lower ability for the ISP to extract surplus from consumers. In fact, our model shows that for low values of κ, consumers might be better off under prioritization than under zero-rating and, eventually, than under net neutrality. Therefore, for severe congestion, a ban on prioritization is likely to lead to ISP choosing zero-rating instead. Our model shows that, in a static perspective, it is the worst-case scenario for consumers but that in a dynamic perspective, it may lead to faster improvement in infrastructure. The welfare effect of a ban on prioritization should trade-off these two dimensions.

Ban on side payments Next, we discuss the consequences of a regulatory policy consisting in banning payments from the CPs to the ISP while taking a laissez-faire approach otherwise, i.e. allowing both prioritization and zero-rating programs without side payments. Whenever side payments do not influence the ISP's regime choice, they are purely transfers from the CPs to the ISP and thus inconsequential for consumer surplus. However, from Proposition 2 we know that side payments can make the ISP switch from a zero-rating regime to a prioritization regime for sufficiently high levels of congestion and advertising revenues. As we have argued above (Corollary 2), it is exactly in those circumstances that consumers could

[22] See p. 37. of the report of the European Commission (2017).

[23] https://tinyurl.com/ycfpaneq

[24] A report by the European Commission (2017) describes the BEREC regulation as well as most of these cases in detail.

[25] https://goo.gl/eHcEJ1

benefit more from prioritization. Therefore, from the consumers' perspective a ban on side-payments can adversely influence the regime choice of the ISP, at least under limited asymmetry between CPs. This is a novel unintended consequence of this policy that our model identifies.

9.2. Conclusion

In the recent public debate about how to regulate the internet, net neutrality advocates have feared that the repeal of regulations would empower ISPs which in turn would harm consumers. On the contrary, opponents of net neutrality rules have claimed that such an empowerment of ISPs is necessary and ultimately beneficial for consumers. Our model reflects this debate by analyzing the two business practices that have recently become unrestricted in the US (prioritization and zero-rating) as new tools for the ISP to discriminate content without any inherent added value. In our baseline model, we focus on the case of content providers that are not too asymmetric. We show that as expected, it is always in the ISP's interest to deviate from net neutrality and use one of the two business practices. In particular, the ISP implements prioritization if congestion is relatively severe and the value of content is high, otherwise it implements a zero-rating plan. Moreover, we find that zero-rating always harms consumers for low levels of congestion. Prioritization is the preferred option of both the ISP and consumers under severe congestion and high-value content, because the low price charged by the ISP to consumers is counterbalanced by large payments from the CPs. Furthermore, we find that the exclusion of a content provider is not optimal for the ISP whenever CPs are not very asymmetric, contrary to the fears of net neutrality advocates.

A first limitation of our model is the monopolistic ISP assumption. We discuss in Section 8 explicitly why we believe that our results are robust to some forms of ISP competition as well. A second limitation is that we consider a model with fixed content supply. In Krämer and Wiewiorra (2012); and Reggiani and Valletti (2016), there is a demand expansion on the consumer side that results from increased participation on the content provider side and different regulations affect content provision by CPs. A third limitation is that our model investigates the ISP's intervention in the content market while assuming that CPs are rather passive. They collect advertising revenues and eventually bid for preferential treatment, but they cannot invest in making their content more attractive to consumers. A natural extension could try modelling investment incentives of CPs, as in Choi et al. (2018). We leave the explicit modeling of these three important features for future work.

CRediT author statement

The two authors have equally contributed to the manuscript.

Appendix A. Extensions and robustness checks

A1. Subscription-based business model

Assume that at the beginning of the game the strong and the weak CP set subscription fees per user f_S and f_W, respectively. Then, the ISP decides on the connection price p and the surcharge ϕ. The location of the indifferent consumer x on the Hotelling segment is:

$$v+q-p-\phi-\tau(x-a)-f_S=v+q-p-\tau(1-x)-f_W \quad \Leftrightarrow \quad x=\frac{1+a}{2}-\frac{\phi}{2\tau}-\frac{f_S-f_W}{2\tau}.$$

Taking their effect on the marginal consumer into account, CPs maximize their total revenue from subscription fees: f_Sx and $f_W(1-x)$. First-order conditions imply that the optimal subscription fees are

$$f_S=\frac{\tau(3+a)-\phi}{3} \quad \text{and} \quad f_W=\frac{\tau(3-a)+\phi}{3}.$$

In this case, the price of CP_S is $\frac{2\tau a}{3}-\frac{2}{3}\phi$ above the price of CP_W and the location of the indifferent consumer is given by $x=\frac{3+a}{6}-\frac{\phi}{6\tau}$. The optimal prices charged by the ISP are:

$$\phi=\frac{a\tau}{2}>0 \quad \text{and} \quad p=v+\kappa-\frac{\tau(6-a)}{4},$$

leading to a profit equal to $\pi_{ISP}^{ZR}=v+\kappa-\frac{3}{2}\tau+\frac{a\tau(12+a)}{24}$. Given that the optimal price differential, ϕ, is strictly positive, this profit is larger than the profit under net neutrality. Therefore despite the positive subscription fees charged by the CPs, the ISP still have some room for zero-rating.

A2. Simultaneous prioritization and zero-rating

In our baseline model, we give to the ISP the choice between two forms of discrimination in price or in quality. In this section, we consider the situation in which the ISP can discriminate both in price *and* in quality, a regime of combined prioritization and zero-rating (P+ZR).

As before, let us write $p_S = p + \phi$, where ϕ can be positive or negative. With priority and zero-rating, the location of the indifferent consumer x is given by:

$$v + q_S - p - \phi - \tau(x - a) = v + q_W - p - \tau(1 - x) \quad \Leftrightarrow \quad x = \frac{1+a}{2} - \frac{\phi}{2\tau} + \frac{q_S - q_W}{2\tau}.$$

Taking qualities (q_S, q_W) as given and solving the ISP's maximization problem, the optimal prices are: $\phi = \frac{a\tau}{2} + \frac{q_S - q_W}{2}$ and $p = v + \frac{q_S + 3q_W}{4} - \frac{\tau}{2} + \frac{a\tau}{4}$. At these prices, the market share of the CP_S is given by $x^{ZR+P} = \frac{2+a}{4} + \frac{q_S - q_W}{4\tau}$. In an interior equilibrium with priority to the strong CP, we have $q_S = 1$ and $q_W = \frac{\kappa - x^{ZR+P}}{1 - x^{ZR+P}}$; with priority to the weak CP, we have $q_S = \frac{\kappa - (1 - x^{ZR+P})}{x^{ZR+P}}$ and $q_W = 1$. Similarly to the prioritization case, finding the solution involves a fixed-point problem.

When the weak CP is prioritized, the prices are:

$$p^{ZR+P} = (1+v) + \frac{a\tau}{4} - \frac{(6+a)\tau}{8} + \frac{1}{8}\sqrt{(2+a)^2\tau^2 - 16(1-\kappa)\tau},$$

$$\phi^{ZR+P} = \frac{1}{4}\sqrt{(2+a)^2\tau^2 - 16(1-\kappa)\tau} - \frac{(2-a)\tau}{4}.$$

And we denote the corresponding profit by π^{ZR+P}. Analogously to Proposition 1, we can show that when combined with zero-rating, the ISP prefers to give priority to the weak content in an interior equilibrium whenever such an equilibrium exists.

Proposition 5. *An interior equilibrium with prioritization and zero-rating exists for* $\kappa \geq 1 - \frac{(2+a)^2\tau}{16}$. *For these parameters and without side payments, the ISP always finds it profitable to prioritize the weak CP. The weak CP is also zero-rated if* $\kappa \geq \kappa_2 = 1 - \frac{a\tau}{2}$, *and the strong CP is zero-rated if* $1 - \frac{(2+a)^2\tau}{16} \leq \kappa \leq \kappa_2$. *However, for the ISP, prioritization combined with zero-rating results in a lower profit than zero-rating alone:* $\pi^{ZR+P} \leq \pi^{ZR}$.

With prioritization of the weak CP, the price differential ϕ is equal to $\frac{a\tau}{2} - \frac{1-q_S}{2}$. It increases with a and decreases with the quality differential. Whenever the capacity is limited, the quality differential is large and the ISP zero-rates the non-prioritized content ($\phi < 0$) and price discrimination mitigates quality discrimination. Conversely, for larger capacity, the quality differential is limited and the prioritized content is also zero-rated. In this case, price discrimination reinforces quality discrimination. However, the profit with prioritization and zero-rating is lower than the profit with zero-rating alone as long as there is some congestion.

Finally, could the combination of priority and zero-rating ever emerge in the presence of side payments? To provide a partial answer to this question, let us consider the following situation. Assume that $\kappa \geq \kappa_2$ and that the ISP chooses zero-rating in equilibrium (see the condition in Proposition 2). The weak content provider benefits from being prioritized on top of being zero-rated as its market share increases compared to the case where it is only zero-rated: $(1 - x^{ZR+P}) \geq (1 - x^{ZR})$. Hence, it is ready to pay more for benefiting from price and quality discrimination. However, the ISP will lose revenue from the consumer side. Then if the benefit of the CP is larger than the loss of the ISP, implementing prioritization and zero-rating against compensation is mutually profitable. Formally, a sufficient condition for this is

$$\rho(x_W^{P+ZR} - x_W^{ZR}) > \pi^{ZR} - \pi^{P+ZR}.$$

This demonstrates that even if the combination of prioritization and zero-rating reduces ISP revenues from the user side, it can emerge at equilibrium if the revenue from the CP side is large enough to compensate it, i.e. if advertising revenue ρ is sufficiently large.

A3. Investment in capacity

In our model, the total welfare (the sum of consumers' surplus and firms' profits) is strictly increasing and linear in capacity level κ. Assuming the cost of capacity building $c(\kappa)$ is strictly increasing and convex, the socially optimal capacity level is given by

$$\min\{1;\ \kappa^o\} \quad \text{where} \quad c'(\kappa^o) = 1.$$

The ISP's overall profit under the zero-rating regime, $\max\{\tilde{p}^{ZR};\ \rho(x^{NN} - x^{ZR}) + p^{NN}\}$, is also linearly increasing in capacity, thus the ISP's investment incentives are completely aligned with the social optimum: it also builds $\min\{1; \kappa^o\}$.

On the contrary, under prioritization, the capacity level chosen by the ISP is generically socially suboptimal. Whenever prioritization is optimal, the ISP's profit equals $p_S^P + \rho(x_S^P - x_W^P)$. The investment in capacity under prioritization is given by

$$\frac{\mathrm{d}p_S^P}{\mathrm{d}\kappa} + \rho\frac{\mathrm{d}(x_S^P - x_W^P)}{\mathrm{d}\kappa} = c'(\kappa)$$

By concavity, the first term is larger (lower) than 1 if κ is lower (larger) than the threshold value κ_5. The second term represents the distortion effect and is thus always negative. Clearly, there exists a cutoff $\widehat{\rho}(\kappa)$ such that the profit-maximizing capacity level is lower than the socially optimal one if either (i) $\kappa \geq \kappa_5$ or (ii) $\rho \geq \widehat{\rho}(\kappa)$.

Proposition 6. *Under zero-rating, the ISP's investment in capacity is always at the socially optimal level of* $\min\{1; \kappa^o\}$. *Under prioritization, the ISP chooses a level of capacity below the socially optimal level whenever (i)* $\kappa \geq \kappa_5$ *or (ii)* $\rho \geq \widehat{\rho}(\kappa)$.

Appendix B. Proofs

B1. Proof of Lemma 1

An interior equilibrium where the ISP gives priority to the weak firm's content solves:

$$v + q_S - p - \tau(x-a) = v + 1 - p - \tau(1-x),$$
$$q_S = 1 - \frac{1-\kappa}{x},$$
$$1 - x \le \kappa.$$

This leads to two potential interior solutions for x:

$$x_1 = \frac{1+a}{4} - \sqrt{\frac{(1+a)^2}{16} - \frac{1-\kappa}{2\tau}} \quad \text{and} \quad x_2 = \frac{1+a}{4} + \sqrt{\frac{(1+a)^2}{16} - \frac{1-\kappa}{2\tau}},$$

which in turn lead to two different prices:

$$p_1 = v + 1 - \tau\left(\frac{3-a}{4} + \sqrt{\frac{(1+a)^2}{16} - \frac{1-\kappa}{2\tau}}\right) < p_2 = v + 1 - \tau\left(\frac{3-a}{4} - \sqrt{\frac{(1+a)^2}{16} - \frac{1-\kappa}{2\tau}}\right).$$

Given that $x_1 < x_2$, it is straightforward that the latter leads to a higher price. Moreover, whenever $1 - x_1 < \kappa$, the condition for the first solution being interior is satisfied, $1 - x_2 < \kappa$ also holds, i.e. the second solution is also interior. Therefore, whenever the first solution is potentially an interior equilibrium, the second solution is also potentially an interior equilibrium and it leads to a higher price for the ISP. Hence the first solution is always dominated by the second and will never constitute an equilibrium of the pricing subgame.

There are two conditions for the existence of an interior solution. The first is that the quadratic equation defining the location of the indifferent consumer have real root(s), which is guaranteed if and only if:

$$\frac{(1+a)^2}{16} - \frac{1-\kappa}{2\tau} \ge 0 \quad \Leftrightarrow \quad \kappa \ge 1 - \frac{\tau(1+a)^2}{8}.$$

The second condition is $1 - x_2 \le \kappa$, which rewrites as

$$\frac{1+a}{4} + \sqrt{\frac{(1+a)^2}{16} - \frac{1-\kappa}{2\tau}} \ge 1 - \kappa \quad \Leftrightarrow \quad \kappa \ge \frac{1+\tau-a\tau}{2\tau}.$$

The second condition implies the first one as $\frac{1+\tau-a\tau}{2\tau} > 1 - \frac{\tau(1+a)^2}{8}$, thus the second condition is necessary and sufficient for the existence of an interior solution of the pricing subgame when the weak CP has priority. To simplify notation in the main text, let $x_W^P \equiv x_2$ and $p^P \equiv p_2$. □

B2. Proof of Proposition 1

In order to prove the Proposition, we show that in all the other equilibrium configurations lead to a lower price than p^P.

Interior equilibrium when CP_S is prioritized If the ISP gives priority to the stronger firm's content, in any interior solution the indifferent consumer's location x is given by

$$v + 1 - p - \tau(x-a) = v + q_W - p - \tau(1-x) \quad \Leftrightarrow \quad x = \frac{1+a}{2} + \frac{1-q_W}{2\tau},$$

where the quality of the weaker CP's content is $q_W = \frac{\kappa - x}{1-x}$. As in the previous case, there are two potential interior solutions for x:

$$x_1' = \frac{3+a}{4} - \sqrt{\frac{(1-a)^2}{16} - \frac{1-\kappa}{2\tau}} \quad \text{and} \quad x_2' = \frac{3+a}{4} + \sqrt{\frac{(1-a)^2}{16} - \frac{1-\kappa}{2\tau}},$$

which in turn lead to two different prices for the ISP:

$$p_1' = v + 1 - \tau\left(\frac{3(1-a)}{4} - \sqrt{\frac{(1-a)^2}{16} - \frac{1-\kappa}{2\tau}}\right) > p_2' = v + 1 - \tau\left(\frac{3(1-a)}{4} + \sqrt{\frac{(1-a)^2}{16} - \frac{1-\kappa}{2\tau}}\right).$$

Analogously to the previous case, it can be shown that one of the solutions leads to a strictly higher price while being always an interior solution whenever the other solution is interior. Therefore, the solution x_2' is dominated, and the condition for the existence of an interior solution is $x_1' < \kappa$ which translates to $\kappa \ge \frac{1+\tau+a\tau}{2\tau}$ which is satisfied when Assumption 4 holds. In the following, we call $p_S^{INT} = p_1'$ and $x_S^P = x_1'$.

We now show that p_S^{INT} is below p^P, i.e.

$$\nu+1-\tau\left(\frac{3(1-a)}{4}-\sqrt{\frac{(1-a)^2}{16}-\frac{1-\kappa}{2\tau}}\right)<\nu+1-\tau\left(\frac{3-a}{4}-\sqrt{\frac{(1+a)^2}{16}-\frac{1-\kappa}{2\tau}}\right).$$

Clearly, this is equivalent to

$$\frac{3(1-a)}{4}-\sqrt{\frac{(1-a)^2}{16}-\frac{1-\kappa}{2\tau}}>\frac{3-a}{4}-\sqrt{\frac{(1+a)^2}{16}-\frac{1-\kappa}{2\tau}}.$$

which rewrites as

$$a/2>\sqrt{\frac{(1+a)^2}{16}-\frac{1-\kappa}{2\tau}}-\sqrt{\frac{(1-a)^2}{16}-\frac{1-\kappa}{2\tau}}.$$

After squaring both sides of the inequality (both sides being positive) and rearranging, we get

$$\frac{(1-a)(1+a)}{16}-\frac{1-\kappa}{2\tau}>\sqrt{\left(\frac{(1+a)^2}{16}-\frac{1-\kappa}{2\tau}\right)\left(\frac{(1-a)^2}{16}-\frac{1-\kappa}{2\tau}\right)}.$$

The price p_S^{INT} is properly defined if $\frac{(1-a)^2}{16}-\frac{1-\kappa}{2\tau}>0$. Hence, both the term on the left-hand side and the term under the square root are positive. Squaring both sides of the inequality and rearranging results in

$$a^2>0 \quad \Leftrightarrow \quad p^P>p_S^{INT}$$

which is what we wanted to show.

Corner equilibrium when CP_S is prioritized A corner equilibrium occurs when the prioritization of the stronger firm makes it so attractive that even the farthest consumer, the one located at 1 prefers buying its product. In this situation, the quality of the prioritized firm is κ, whereas the quality of its competitor is 0. The strong firm can then extract all the surplus of the farthest customer and thus charges $p_S^C=\nu+\kappa-\tau(1-a)$. We must then show that

$$p_S^C<p^P \quad \Leftrightarrow \quad \nu+\kappa-\tau(1-a)<\nu+1-\tau\left(\frac{3-a}{4}-\sqrt{\frac{(1+a)^2}{16}-\frac{1-\kappa}{2\tau}}\right)$$

which after rearranging the terms rewrites as

$$\sqrt{\frac{(1+a)^2}{16}-\frac{1-\kappa}{2\tau}}>\frac{3a-1}{4}-\frac{1-\kappa}{\tau}.$$

Notice that the left-hand side is always positive while the right-hand side is always negative by the assumptions $a<2/7$ and $\kappa<1$. Therefore $p_S^C<p^P$ always holds.

Corner equilibrium when CP_W is prioritized A corner equilibrium occurs when the prioritization of the weak CP makes it so attractive that even the farthest consumer, the one located at 0 prefers buying from it. In this situation, the quality of the prioritized firm is κ, whereas the quality of its competitor is 0. The weak firm can then extract all the surplus of the farthest customer and thus charges $p_W^C=\nu+\kappa-\tau$. Notice that $p_W^C=\nu+\kappa-\tau<p_S^C=\nu+\kappa-\tau(1-a)$, and above we showed that $p_S^C<p^P$, therefore $p_W^C<p^P$ always holds.

Semi-corner equilibrium when CP_S is prioritized A semi-corner equilibrium occurs when the non-prioritized (weaker) firm provides 0 quality but still serves some consumers thanks to the large value of internet connection ν. Such a situation arises whenever the indifferent consumer is such that $x\geq\kappa$, leading to the stronger firm providing quality κ/x. The location of the indifferent consumer x is given by

$$\nu+\frac{\kappa}{x}-p-\tau(x-a)=\nu+0-p-\tau(1-x),$$

which leads to two potential solutions:

$$\frac{1+a}{4}-\sqrt{\frac{(1+a)^2}{16}+\frac{\kappa}{2\tau}} \quad \text{and} \quad x_S^{SC}=\frac{1+a}{4}+\sqrt{\frac{(1+a)^2}{16}+\frac{\kappa}{2\tau}}.$$

However, it is easy to see that the first one is negative, thus we can focus our attention to x_S^{SC}, leading to the price of

$$p_S^{SC}=\nu-\tau(1-x_S^{SC})=\nu-\tau\left(\frac{3-a}{4}-\sqrt{\frac{(1+a)^2}{16}+\frac{\kappa}{2\tau}}\right)$$

therefore

$$p_S^{SC}<p^P \quad \Leftrightarrow \quad \sqrt{\frac{(1+a)^2}{16}+\frac{\kappa}{2\tau}}-\sqrt{\frac{(1+a)^2}{16}-\frac{1-\kappa}{2\tau}}<\frac{1}{\tau}.$$

Both sides of the inequality being positive, squaring them and rearranging leads to

$$L \equiv \frac{(1+a)^2}{16} + \frac{\kappa}{2\tau} - \frac{1}{4\tau} - \frac{1}{2\tau^2} < \sqrt{\frac{(1+a)^2}{16} + \frac{\kappa}{2\tau}}\sqrt{\frac{(1+a)^2}{16} - \frac{1-\kappa}{2\tau}}$$

This inequality is satisfied if either (i) the term on left-hand side is non-positive, i.e. $L \le 0$; or (ii) the square of both positive numbers also satisfy the inequality. (i) can be rewritten as

$$L \le 0 \quad \Leftrightarrow \quad \kappa \le \frac{4\tau - a^2\tau^2 - 2a\tau^2}{8\tau} + \frac{8-\tau^2}{8\tau} \equiv \kappa',$$

whereas (ii) leads to

$$\kappa > \frac{4\tau - a^2\tau^2 - 2a\tau^2}{8\tau} + \frac{4}{8\tau} \equiv \kappa''.$$

For $\tau \le 2$ we have $\kappa'' \le \kappa'$ thus (i) or (ii) is satisfied for any κ, therefore $p_S^{SC} < p^P$.

For $\tau > 2$, it is easy to check that $\kappa^* = \frac{1+\tau-a\tau}{2\tau} > \kappa''$ thus Assumption 4 guarantees that (ii) is always satisfied, leading to $p_S^{SC} < p^P$.

Semi-corner equilibrium when CP_W is prioritized Finally, we investigate the equilibria leading to a situation where the non-prioritized (stronger) firm provides 0 quality while still serving some consumers due to the large value of internet connection v. Such a situation arises whenever the indifferent consumer (at position x) is located in the $[0, 1-\kappa)$ interval. In a semi-corner equilibrium the weaker firm provides quality $\kappa/(1-x)$ to the $1-x$ consumers closest to it.

We have to distinguish two cases depending on the location of the indifferent consumer. Firstly, if it is located to the right of CP_S, i.e. $x \in (a, 1-\kappa)$, then x is given by

$$v + 0 - p - \tau(x-a) = v + \frac{\kappa}{1-x} - p - \tau(1-x),$$

which leads to two potential solutions:

$$x_1 = \frac{3+a}{4} - \sqrt{\frac{(1-a)^2}{16} + \frac{\kappa}{2\tau}} \quad \text{and} \quad x_2 = \frac{3+a}{4} + \sqrt{\frac{(1-a)^2}{16} + \frac{\kappa}{2\tau}}.$$

However, it is easy to show that $x_2 > 1$, thus we can focus our attention to x_1. Notice that one of the conditions for x_1 being a semi-corner equilibrium, $x_1 < 1-\kappa$, rewrites as

$$\kappa - \frac{1-a}{4} < \sqrt{\frac{(1-a)^2}{16} + \frac{\kappa}{2\tau}}.$$

Assumption 4 implies $\kappa > \frac{1-a}{2}$ thus both sides of the inequality are positive. Squaring and rearranging reveals that $\kappa < \frac{1+\tau-a\tau}{2\tau}$ is a necessary condition for the existence of such an equilibrium, which is ruled out by Assumption 4.

Secondly, if the indifferent consumer is located to the left of CP_S, i.e. $x \in (0, a]$, then its location is given by

$$v + 0 - p - \tau(a-x) = v + \frac{\kappa}{1-x} - p - \tau(1-x),$$

which leads to

$$x_W^{SC3} = 1 - \frac{\kappa}{\tau(1-a)} \quad \text{and} \quad p_W^{SC3} = v - \frac{\kappa}{1-a} + \tau(1-a).$$

Notice that one of the necessary conditions for the existence of such a semi-corner solution is $x_W^{SC3} < a \quad \Leftrightarrow \quad \tau(1-a)^2 < \kappa$, therefore $\tau(1-a)^2 < 1$ is also necessary.

Next, notice that the price p_W^{SC3} is decreasing in κ whereas the price p^P is increasing in κ. We will prove $p_W^{SC3} < p^P$ by showing that it is satisfied even at the lower bound of possible values of κ. To see this, replacing $\kappa = \kappa^*$ into the prices we have

$$p_W^{SC3} < p^P \quad \Leftrightarrow \quad \tau\left(\sqrt{\frac{(a\tau+\tau-2)^2}{\tau^2}} + 5a - 7\right) + \frac{2}{\tau(1-a)} + 6 > 0.$$

We distinguish two cases. For $\tau(1+a) < 2$, the second inequality simplifies to $(a-2)\tau + \frac{1}{2\tau-2a\tau} + 2 > 0$. It is straightforward to show that the expression on the left-hand side is strictly increasing in a and strictly decreasing in τ. Therefore, it is always larger than its value at $a = 0$ and $\tau = \frac{1}{(1-a)^2}$ (from the existence condition), which equals 1/2, so the inequality is always satisfied. Next, for $\tau(1+a) > 2$ the inequality $p_W^{SC3} < p^P$ rewrites as $-\frac{3}{2}(1-a)\tau + \frac{1}{2\tau(1-a)} + 1 > 0$. Similarly to the previous case, the expression on the left-hand side is strictly increasing in a and strictly decreasing in τ, thus, it is always larger than its value at $a = 0$ and $\tau = \frac{1}{(1-a)^2}$ which is exactly 0.

This concludes the proof of Proposition 1. □

B3. Proof of Proposition 2

The proof consists of two steps. As a first step, the next Lemma states that whenever the interior equilibrium with priority to CP_S exists, it dominates all the other equilibrium with priority given to CP_S.

Lemma 5. *For $\kappa \geq \frac{1+\tau+a\tau}{2\tau}$ an interior equilibrium where the ISP gives priority to the strong content exists and it dominates all other potential equilibria of the pricing subgames involving prioritization of the strong CP. The ISP's profit is then given by*

$$p_S^{INT} = v + 1 - \tau\left(\frac{3(1-a)}{4} - \sqrt{\frac{(1-a)^2}{16} - \frac{1-\kappa}{2\tau}}\right)$$

.

For notational simplicity in the main text, we use $p_S^P = p_S^{INT}$. The result follows from straightforward calculations showing that under the condition of the Lemma, we have

$p_S^{INT} > p_S^{SC}$ and $p_S^{INT} > p_S^C$. Therefore the interior equilibrium is best equilibrium for the ISP whenever it gives priority to CP_S, as stated in the main text.

As a second step, notice that for sufficiently large values of ρ, irrespectively of the value of κ, the profit of the ISP under prioritization is $\rho(x_S^P - x_W^P) + p_S^P$, whereas its profit under zero-rating is $\rho(x^{NN} - x^{ZR}) + p^{NN}$. All of the terms being positive, by the Archimedean property of the real numbers, the former expression is larger for large ρ and prioritization is chosen if and only if

$$x_S^P - x_W^P > x^{NN} - x^{ZR}.$$

It is straightforward to check that the right-hand side is positive and independent of κ, whereas left-hand side is decreasing in κ, moreover, it goes to zero as κ goes to one. By continuity, for values of κ close to one, zero-rating is the preferred option. Next we show that for $\kappa \leq \kappa_1$, the left-hand side is larger. Indeed, $x_S^P > x^{NN}$ always holds. Moreover, for $\kappa \leq \kappa_1$ Lemma 2 implies $x_W^P < x^{ZR}$. Thus, for $\kappa = \kappa_1$, there always exists $\overline{\rho}(\kappa) > 0$ such that prioritization is the equilibrium choice of the ISP. Finally, by continuity there must exist a cut-off value κ' in the $(\kappa_1, 1)$ interval such that for all $\kappa \leq \kappa'$ the ISP chooses prioritization, and it chooses zero-rating otherwise. □

B5. Proof of Proposition 4

For low levels of b, the market will still be fully covered and the maximization problem of the ISP is the following:

$$\max_{p,\phi}\ (1+b)p + (\tilde{x} - (1-b))\phi \quad \text{s.t.} \quad U(\tilde{x}) \geq 0 \quad \text{and} \quad U(1-b) \geq 0.$$

It is straightforward to derive $\tilde{x} = \frac{(3+a)\tau - \phi}{2\tau}$. The first constraint then rewrites as:

$$U(\tilde{x}) \geq 0 \Leftrightarrow v + \frac{\kappa}{(1+b)} - p - \phi - \tau(\tilde{x} - 1 - a) \geq 0 \Leftrightarrow p + \phi/2 \leq v + \frac{\kappa}{(1+b)} - \frac{(1-a)\tau}{2}.$$

Whereas the second constraint simplifies to:

$$U(1-b) \geq 0 \Leftrightarrow v + \frac{\kappa}{(1+b)} - p - \phi - \tau(a+b) \geq 0 \Leftrightarrow p + \phi \leq v + \frac{\kappa}{(1+b)} - \tau(a+b).$$

Clearly, as the profit is increasing in p, at least one of the two constraints must be binding at the optimal solution. We distinguish three regions depending on the size of b.

For small b, in optimum we have $U(\tilde{x}) = 0$ and $U(1-b) > 0$, which is a direct extension of the baseline model (i.e. for $b = 0$). In this region the ISP optimally zero-rates the weak CP, as before, with $\phi^* = \frac{(a+b)\tau}{2} > 0$ increasing in b. This solution applies for $b \in [0, b_0]$, with $b_0 = \frac{2-7a}{5}$. This cut-off value of b is the solution of $U(1-b) = 0$ at the optimal prices.

For medium b, having both constraints binding is optimal: $U(\tilde{x}) = 0$ and $U(1-b) = 0$ are jointly satisfied, leading to the optimal additional fee of $\phi^* = (1 - 3a - 2b)\tau$, decreasing in b. Notice that this optimal additional fee charged to the strong CP is positive first, but strictly decreasing in b, and turns negative at $b = b_1 = (1-3a)/2$. Therefore, in this region as additional demand increases, the ISP first optimally zero-rates the weak ($b \leq b_1$), then the strong CP ($b \geq b_1$).

For the ISP, there is a cut-off value b_2 above which it is no longer optimal to have a fully covered market i.e. for $b \geq b_2$, we will have $U(1-b) \leq 0$ and $\underline{x} \geq 1-b$. The value b_2 is found by replacing the binding constraint $U(1-b) = 0$ by the binding constraint $U(\underline{x}) = 0$ in the above maximization program.[26] The cut-off value b_2 is equal to $b_2 = \frac{v+\tau(2-8a)}{6\tau}$ and the above solution applies for $b \in [b_0, b_2]$. It is easy to check that $b_0 < b_1 < b_2$ under the assumption that $v \geq \tau$.

For large b, namely for $b > b_2$ the consumer located at $1-b$ is not willing to buy. In this case the marginal consumer of the weak CP $\underline{x}$ is defined by $\underline{x} = 1 - b_2$ and the ISP optimally zero rates the strong content applying $\phi^* = (1 - 3a - 2b_2)\tau < 0$.

[26] Or alternatively by finding the profit maximizing value of b in the covered market case i.e. the market size that maximizes the ISP's profit at optimal prices.

Finally, we show that when firms bid to be zero-rated, the ISP will always implement the optimal zero-rating program described above. To show that, we need to show that when CPs compete to be zero-rated, the strong CP cannot overbid the weak one when $b \le b_1$ and, conversely that the weak cannot overbid the strong when $b \ge b_1$.

As before, the alternative to the optimal zero-rating (ϕ^*) corresponds to the neutral situation where $\phi^* = 0$ and the price p extracts all the surplus of the marginal consumer, which is either $\tilde{x} \in [1+a, 2]$ or $\underline{x}$.

When $b \le b_1$, the price is given by the binding constraint $U(\tilde{x}) = 0$ and the market is fully covered under both NN and ZR. Hence, as before the maximal bid of the two firms are equal and the ISP implements the optimal ZR program.

When $b \ge b_1$, the price under NN is set to extract all the surplus of the $\underline{x}$ consumer. It is easy to show that when $\phi^* = (1-3a-2b)\tau < 0$, the indifferent consumer $\tilde{x}$ is the same under ZR than under NN. Therefore, the weak CP will not participate in the zero-rating auction and the ISP zero rates the strong CP. □

B4. Proof of Proposition 5

In an equilibrium with priority to the strong CP, we have:

$$p = v + 1 - \frac{5(2-a)\tau}{8} + \frac{3}{8}\sqrt{\tau}\sqrt{(2-a)^2\tau - 16(1-\kappa)}, \phi = \frac{(2+a)\tau}{4} - \frac{\sqrt{\tau}}{4}\sqrt{(2-a)^2\tau - 16(1-\kappa)} > 0,$$

$$x_S^{P+ZR} = \frac{1}{8\tau}\left(6\tau + a\tau - \sqrt{\tau}\sqrt{(2-a)^2\tau - 16(1-\kappa)}\right),$$

$$\pi_S^{P+ZR} = v + \frac{1+\kappa}{2} + \frac{1}{16}\left((2-a)\sqrt{\tau}\sqrt{(2-a)^2\tau - 16(1-\kappa)} - \tau(2-a)(6+a)\right).$$

An interior equilibrium with priority to the strong CP exists when $\kappa \ge 1 - \frac{(2-a)^2\tau}{16}$.

In an equilibrium with priority to the weak CP, we have:

$$p = v + 1 - \frac{(6-a)\tau}{8} + \frac{1}{8}\sqrt{\tau}\sqrt{(2+a)^2\tau - 16(1-\kappa)}, \phi = \frac{1}{4}\left(-\tau(2-a) + \sqrt{\tau}\sqrt{(2+a)^2\tau - 16(1-\kappa)}\right),$$

$$x_W^{P+ZR} = \frac{1}{8\tau}\left(\tau(2+a) + \sqrt{\tau}\sqrt{(2+a)^2\tau - 16(1-\kappa)}\right),$$

$$\pi_W^{P+ZR} = v + \frac{1+\kappa}{2} + \frac{1}{16}\left((2+a)\sqrt{\tau}\sqrt{(2+a)^2\tau - 16(1-\kappa)} - \tau(2-a)(6+a)\right).$$

An interior equilibrium with priority to the weak CP exists when $\kappa \ge 1 - \frac{(2+a)^2\tau}{16}$, a softer condition than for the existence of the equilibrium with priority to the strong CP.

It is straightforward to show that $\pi_W^{P+ZR} > \pi_S^{P+ZR}$. When the weak CP has priority, it is also zero-rated if $\phi > 0$, which is equivalent to: $\kappa \ge 1 - \frac{a\tau}{2}$. This cut-off value is larger than the existence condition. Therefore, there is an non-empty parameter space $\kappa \in [1 - \frac{(2+a)^2\tau}{16}, 1 - \frac{a\tau}{2}]$ where the strong CP is zero-rated, corresponding to $\phi < 0$.

Finally, we show that $\pi_W^{P+ZR} \le \pi^{ZR}$. After simplifications, this inequality is equivalent to

$$\tau(2+a)^2 - 8(1-\kappa) \ge (2+a)\sqrt{\tau}\sqrt{(2+a)^2\tau - 16(1-\kappa)}.$$

Squaring both sides and simplifying, by Assumption 4 we have $\pi_W^{P+ZR} \le \pi^{ZR}$. □

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.ijindorg.2020.102662.

References

Armstrong, M., 2006. Competition in two-sided markets. Rand J Econ 37 (3), 668–691.
Armstrong, M., Wright, J., 2007. Two-sided markets, competitive bottlenecks and exclusive contracts. Economic Theory 32 (2), 353–380.
BEREC, 2016. Report on the outcome of the public consultation on draft berec guidelines on the implementation by national regulators of european net neutrality rules. BoR 16 (128).
Bourreau, M., Kourandi, F., Valletti, T., 2015. Net neutrality with competing internet platforms. J Ind Econ 63 (1), 30–73.
Broos, S., Gautier, A., 2017. The exclusion of competing one-way essential complements: implications for net neutrality. Int. J. Ind Organiz 52, 358–392. doi:10.1016/j.ijindorg.2017.03.003.
Cheng, H.K., Bandyopadhyay, S., Guo, H., 2011. The debate on net neutrality: a policy perspective. Information systems research 22 (1), 60–82.
Cho, S., Qiu, L., Bandyopadhyay, S., 2016. Should online content providers be allowed to subsidize content? - an economic analysis. Information Systems Research 27 (3), 580–595.
Choi, J.P., Jeon, D.-S., Kim, B.-C., 2015. Net neutrality, business models, and internet interconnection. American Economic Journal: Microeconomics 7 (3), 104–141.
Choi, J.P., Jeon, D.-S., Kim, B.-C., 2018. Net neutrality, network capacity, and innovation at the edges. J Ind Econ 66 (1), 172–204.
Choi, J.P., Kim, B.-C., 2010. Net neutrality and investment incentives. Rand J Econ 41 (3), 446–471.

Clarke, R.N., 2014. Expanding mobile wireless capacity: the challenges presented by technology and economics. Telecomm Policy 38 (8), 693–708. doi:10.1016/j.telpol.2013.11.006.
Easley, R.F., Guo, H., Krmer, J., 2018. Research commentary-from net neutrality to data neutrality: a techno-economic framework and research agenda. Information Systems Research 29 (2), 253–272. doi:10.1287/isre.2017.0740.
Economides, N., 1989. Quality variations and maximal variety differentiation. Reg Sci Urban Econ 19 (1), 21–29.
Economides, N., Hermalin, B.E., 2012. The economics of network neutrality. Rand J Econ 43 (4), 602–629.
Economides, N., Tåg, J., 2012. Network neutrality on the internet: a two-sided market analysis. Information Economics and Policy 24 (2), 91–104.
European Commission, 2017. Zero-rating practices in broadband markets. Luxembourg: Publications Office of the European Union.
Greenstein, S., Peitz, M., Valletti, T., 2016. Net neutrality: a fast lane to understanding the trade-offs. The Journal of Economic Perspectives 30 (2), 127–149.
Heikkinen, M.V., Berger, A.W., 2012. Comparison of user traffic characteristics on mobile-access versus fixed-access networks. In: International Conference on Passive and Active Network Measurement. Springer, pp. 32–41.
Hoernig, S., Monteiro, F., 2018. Zero-rating, network effects, and capacity investments. mimeo.
Inceoglu, F., Liu, X., 2019. Multiproduct price discrimination with quantity limits: an application to zero-rating. Econ Lett 180, 41–45.
Inderst, R., Peitz, M., 2014. Investment under uncertainty and regulation of new access networks. Information Economics and Policy 26, 28–41. doi:10.1016/j.infoecopol.2013.11.001.
Jeitschko, T.D., Kim, S.J., Yankelevich, A., 2017. Zero-rating and vertical content foreclosure. mimeo.
Jeitschko, T.D., Kim, S.J., Yankelevich, A., 2018. A cautionary note on using hotelling models in platform markets. DICE Discussion Papers. University of Dsseldorf, Dsseldorf Institute for Competition Economics (DICE).
Jullien, B., Sand-Zantman, W., 2018. Internet regulation, two-sided pricing, and sponsored data. Int. J. Ind Organiz 58, 31–62. doi:10.1016/j.ijindorg.2018.02.007.
Krämer, J., Peitz, M., 2018. A fresh look at zero-rating. Telecomm Policy 42 (7), 501–513. doi:10.1016/j.telpol.2018.06.005.
Krämer, J., Wiewiorra, L., 2012. Network neutrality and congestion sensitive content providers: implications for content variety, broadband investment, and regulation. Information Systems Research 23 (4), 1303–1321.
Krämer, J., Wiewiorra, L., Weinhardt, C., 2013. Net neutrality: a progress report. Telecomm Policy 37 (9), 794–813.
Malone, J.B., Nevo, A., Williams, J.W., 2017. The tragedy of the last mile: congestion externalities in broadband networks. NET Institute Working Paper No. 16–20..
Marsden, C.T., 2010. Net neutrality: Towards a co-regulatory solution. A&C Black.
Neven, D., Thisse, J.-F., 1989. On Quality And Variety Competition. CORE Discussion Papers. Universit catholique de Louvain, Center for Operations Research and Econometrics (CORE).
Nevo, A., Turner, J.L., Williams, J.W., 2016. Usage-based pricing and demand for residential broadband. Econometrica 84 (2), 411–443.
Peitz, M., Schuett, F., 2016. Net neutrality and inflation of traffic. Int. J. Ind Organiz 46, 16–62. doi:10.1016/j.ijindorg.2016.03.003.
Reggiani, C., Valletti, T., 2016. Net neutrality and innovation at the core and at the edge. Int. J. Ind Organiz 45, 16–27.
Schnurr, D., 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. International Telecommunications Society (ITS).
Somogyi, R., 2018. Zero-Rating and Net Neutrality. Revised version of CORE Discussion Papers. Université catholique de Louvain.
Van Schewick, B., 2016. T-Mobile's binge on violates key net neutrality principles. Stanford Law School's Center for Internet and Society report.
Yoo, C.S., 2016. Avoiding the pitfalls of net uniformity: zero rating and nondiscrimination. Review of Industrial Organization 1–28.