Working Paper
627

2018

Zero-rating, network effects, and capacity investments

Steffen Hoernig
Francisco Monteiro

Zero-Rating, Network Effects, and Capacity Investments

Steffen Hoernig
Nova School of Business and Economics
Lisbon, Portugal*

Francisco Monteiro
Nova School of Business and Economics
Lisbon, Portugal†

July 9, 2018

Abstract

We consider internet service providers' incentives to zero-rate, i.e. do not count towards data allowances, the consumption of certain services, in the absence of payments from content providers. In a general model with various types of network effects, service substitutes or complements, monopoly and duopoly, we show that ISPs adopt zero-rating and that it increases consumer surplus and total welfare if network effects are strong enough. Capacity investment increases (decreases) with network effects if services are complements (substitutes). Under competition, the decision to zero-rate depends the residual network effect, which includes the impacts of spillovers and brand differentiation.

JEL Classification: D21; L51; L96.

Keywords: Zero-rating; Network effects; Net neutrality; Capacity Investment.

*Universidade Nova de Lisboa, Campus de Campolide, 1099-032 Lisboa, Portugal, shoernig@novasbe.pt. This work was funded by National Funds through FCT – Fundação para a Ciência e Tecnologia under the project Ref. UID/ECO/00124/2013 and by POR Lisboa under the project LISBOA-01-0145-FEDER-007722.

†Universidade Nova de Lisboa, Campus de Campolide, 1099-032 Lisboa, Portugal, 22040@novasbe.pt. This paper arose from Francisco' MSc Thesis at Nova SBE.

1 Introduction

1.1 Zero-Rating

The term "zero-rating" refers to the recent practice of internet service providers (ISPs) to exempt certain kind of data traffic from data caps under fixed or mobile broadband subscriptions. That is, subscribers can access these services while the corresponding data consumption is not deducted from their monthly data allowance. Services that are often zero-rated are social networks (such as Facebook, Whatsapp, Instagram, Twitter) and video services (Netflix). This practice is controversial, as it may conflict with the principle of "net neutrality" which states that all internet traffic should be treated equally.

The principle of net neutrality was implemented in the U.S. through the "2015 Open Internet Order" (FCC 2015). This order banned providing a "fast lane" against payment by content providers (CPs), but left zero-rating arrangements to a case-by-case review due to potential consumer benefits (p.66).[1] The European Union created its “EU Open Internet Access Regulation 2015” (Regulation (EU) 2015/2120), in force since 30 April 2016 (EC 2015); it does not mention zero-rating explicitly. The net neutrality Guidelines (BEREC 2016) of the association of European national regulatory agencies provide for a differentiated treatment of zero-rating (p. 11, para. 40-43). Offers which block or slow down all traffic except the zero-rated services once the data limit is reached are always considered in violation of net neutrality rules, whereas offers that zero-rate specific a service, rather than a class of competing services, are more likely to be considered as such. Any assessment needs to take into account the principles on which the Open Internet Regulation is based.

The market has been extensively monitored: by the FCC (2017), BEREC (2017), Ofcom (2017), the European Commission (EC 2017). Some national regulators have prohibited some or all zero-rating tariffs. In the Netherlands, the Telecommunications Act of May 2016 included a blanket prohibition of price discrimination including zero-rating (This was struck down by a court in April 2017). The German Bundesnetzagentur[2] in December 2017 and the Portuguese ANACOM[3] in March 2018 declared certain tariffs to be in violation of the Open Internet and Roaming Regulations.

1.2 Our Contribution

While we are well aware of the concerns voiced in the Regulations and studies mentioned above, in this paper we set out to explore further the consumer and social benefit side of zero-rating. In Europe payments from CPs to ISPs are illegal under the present net neutrality regulations, but still we see zero-rating offers proliferating. Thus other factors must make zero-rating attractive to ISPs. Therefore, and contrary to the previous literature,

---

[1] The 2015 FCC order of the Obama era was overturned in 2018 under the Trump administration (FCC 2018).

[2] See https://www.bundesnetzagentur.de/SharedDocs/Pressemitteilungen/DE/2017/15122017_StreamOn.html?nn=473132.

[3] See https://www.anacom.pt/render.jsp?contentId=1430837&languageId=1.

we exclude payments from content providers to ISPs as the reason for the zero-rating of certain services. Instead, we explore whether network effects can be a sufficiently strong motive for zero-rating, and whether consumers and society benefit from this.[4] As we show below, zero-rating in this context has the role of allocating sparse network capacity efficiently, taking into account bandwidth usage and the aggregate value of content consumption.

We build a model with different types of network effects (club effects, firm-level effects or market level spillovers), monopoly or duopoly ISPs with general specifications of subscription and usage (substitutes or complements) demands, capacity constraints and capacity investment. ISPs endogenously decide whether to adopt zero-rating or not. While the treatment is general, it is illustrated with a specific model in the Appendix that satisfies the assumptions made in the text. We model zero-rating as a choice on a continuum between full zero-rating (one service is counted not at all towards the data cap) and a joint data cap (both services are fully counted). This allows us to derive precise optimality conditions.

Our main results are as follows, first for a monopoly ISP. Zero-rating is indeed a profit-maximizing choice for the ISP if network effects are strong enough, and even more so if costs of increasing network capacity are low. Here the ISP shifts capacity usage towards the service that creates network effects in order to increase total surplus, independently of whether services are substitutes or complements. Still, the latter matters for investment: under zero-rating, stronger network effects decrease capacity investment if services are substitutes, and increase it if they are complements.

Considering only producer and consumer surplus, if the ISP chooses zero-rating then it is socially optimal to do so, both under club and firm-level network effects. Taking into account externalities on third parties only reverses this outcome if the externalities resulting created by the usage of the non-zero-rated service are very strong. The ISP's capacity investment decision under zero-rating is socially optimal either club effects, if services are independent, or if the whole market is covered. Otherwise the ISP overinvests if services are substitutes and underinvests if they are complements.

Under duopoly competition between ISPs, we find that the decisive factor for zero-rating and capacity decisions is the "residual network effect". It captures three forces: (club or firm-level) network effects on each ISP's own subscriber base, the spillover of network effects to subscribers of the other ISP, and brand differentiation. The latter measures the displacement of customers from the ISP to its rival when surplus changes. Zero-rating remains profit-maximizing under competition if the residual network effect is strong enough (in particular, of course, if there are no spillovers to start with). We also find that ISPs with a larger customer base are more prone to introduce zero-rating based on network effects.

[4] In a working paper for NERA, Eisenach (2015) stresses how zero-rating under network effects can increase market participation, in particular in developing countries. He does not consider the issue of whether network effects make operators adopt zero-rating in the first place.

1.3 Relation to the Literature

Zero-rating has been the subject of little attention in the academic literature so far.[5] Yoo (2017), arguing from a legal point of view, defends that zero-rating should be allowed because it allows ISPs to differentiate their tariff offers. Kramer and Peitz (2018) provide a policy-focused discussion of zero-rating, weighing the benefits and potential social costs of different implementations. They give particular attention to "throttling", i.e. reductions in transmission quality, of the zero-rated services. Similar to our paper, they consider reasons why ISPs might adopt zero-rating in the absence of payments from CPs, arguing that arrangements involving payments are illegal under the existing European net neutrality rules.[6]

Other papers deal explicitly with payments from CPs to ISPs. Jullien and Sand-Zantman (2017) consider zero-rating as an instrument to price discriminate between CPs that provide services that are independent in consumption but of different value to them. The ISP charges them for "sponsored data", i.e. zero-rated traffic, and sets a data cap that restricts usage to the amount where the marginal benefits to the ISP does not exceed its marginal costs. Sponsored data plans are taken up by high-value CPs (for example, those with high revenues from advertising) and allow the ISP to bring the consumption of high-value content to the efficient level. In the presence of zero rating, this data cap is lowered strategically to reduce consumption of the non-zero-rated content, which increases the ISP's profits from the zero-rated operator. Jullien and Sand-Zantman do not explicitly consider capacity investments, even though they include a long-run marginal cost of usage. We show in our paper that zero rating can arise in the absence of payments by CPs, and also take account of the substitutability between contents. The latter turns out to be important to gauge the impact of network effects on investment incentives. Furthermore, contrary to their model, in our setting the zero-rating decision is affected by competition, through potential spillovers of network effects.

Somogyi (2017) considers the choice between "exclusive zero-rating", where only specific services are zero-rated, and "open zero-rating", where whole classes of services are zero-rated, again in the context of payments by CPs to the ISP. He explicitly takes into account the ISP's capacity constraint and assumes that consumers are rationed at the subscription (rather than the usage) stage. The retail tariff is given exogenously, which affects the computation of profits in the different scenarios he considers. We model a capacity constraint which the ISP takes into account when setting the (non-zero-rated) data cap and the retail tariff. The latter is endogenous and depends on the zero-rating scenario chosen by the ISP.

Schnurr and Wiewiorra (2018) consider both zero-rating (without payments from CPs) and sponsored data as devices to support price discrimination via data caps between different consumer types. Zero-rating of services which consumers value similarly increases effective heterogeneity and the degree of rent extraction for services where valuation is different. The authors show that zero-rating in this context reduces consumer surplus,

[5] It has received quite more attention in regulatory circles, though. There was also a draft paper by Inceoglu and Liu, which at the time of this writing is no longer publicly available.

[6] This does not mean that Kramer and Peitz support those rules. Rather, they argue that these rules impede the efficient functioning of the market due to their blanket nature.

while sponsored data may increase it if choice is maintained. Contrary to our paper, they assume that network capacity is unlimited, so that zero-rating does not have the role of allocating scarce capacity more efficiently. Rather, it is shown that under some circumstances zero-rating leads to higher profits than a uniform data cap.[7]

Jaunaux and Lebourges (2018) provide a legal and economic overview, from an operator perspective, of the short- and longer-run effects of zero-rating on consumers and the provision of content. They conclude that in the short term zero-rating benefits consumers, and that also in the longer run their freedom of choice is not restricted if sponsored data plans are open rather than closed. Problems would only arise if ISP and CP are dominant, so that competitive forces are not strong enough to maintain choice.

Our paper is related to the already vast literature on net neutrality from a general point of view. Here we only refer to the recent overview by Greenstein, Peitz and Valletti (2016). They point out in particular that no simple general statements can be made about the impact of net neutrality rules on investment incentives.

Jeon, Laffont and Tirole (2004) analyze interconnection between voice telephony operators in the presence of receiver benefits. They find that operators strategically set higher retail prices for calls between networks in order to curtail the receiver benefits on rival networks.[8] These receiver benefits are similar to the competitive spillover of network effects in our model, and ISPs equally have an incentive to limit the benefits that competitors' subscribers obtain. The strategic effect of spillovers in our setting is that they make zero-rating less attractive. Schmutzler (2013) considers spillovers in cost-reducing investments and shows that these may increase or decrease equilibrium investments under competition. In our setting network effect spillovers always decrease investment incentives.

2 A Model of Zero-Rating with Network Effects

2.1 Setup

In this section we consider the case of a monopolist ISP in general terms. In the appendix we also provide a specific model of usage and subscription demands which satisfies all the assumptions made in the text and is used below for some numerical simulations.

Consumers obtain a utility of $S(x, y, z) = U(x, y) + \beta z$ from two services x and y provided by the ISP plus a network effect from $z = nx$, the aggregate consumption of service x.[9] U is twice continuously differentiable and increases in both services ($U_x, U_y > 0$),[10] for service x up to bliss points $\bar{x}(y)$ such that $U_x(\bar{x}(y), y) = 0$, and is strictly concave in (x, y), i.e. $U_{xx}, U_{yy} < 0$ and $U_{xx}U_{yy} - U_{xy}^2 > 0$. Services are substitutes if $U_{xy} < 0$ and complements if $U_{xy} > 0$.

We assume that network effects raise utility, $\beta > 0$. Different models of network effects are encapsulated in our formulation. "Club effects", i.e. network effects limited

[7] It can be shown that in their model zero-rating always leads to lower profits than separate caps for different types of content, thus they never are an optimal choice of ISPs.

[8] See also Hoernig (2007) and Hoernig (2014).

[9] The values of x and y can be interpreted as time spent on each of the two services.

[10] We will use subscripts for partial derivatives.

to specific groups of users can be captured by $n = v$ with constant $v > 0$. "Firm-level network effects", which depend on network size α, are described by $n = \alpha$. Below we expand the latter definition to market-level spillovers. The assumption that utility from usage and network effects is additively separable implies that consumers' usage choices do not depend on z: Network effects only count for subscription decisions.

While consumers are homogeneous with respect to usage, they are heterogeneous with respect to the benefits derived directly from subscribing to the ISP. The number of subscribers α is given by the differentiable and strictly increasing function G of the net benefits from usage minus payments. The monopolist ISP obtains all content at zero cost, but has a cost c per unit of capacity Q.

We model zero-rating as follows: Service x consumes $r > 0$ units and service y one unit of capacity per unit of usage. The ISP counts a share $\lambda \in [0, 1]$ of the capacity usage of service x against the cap, with $\lambda = 0$ if it zero-rates x and $\lambda = 1$ if both services count towards the cap.[11] It offers a flat-rate tariff (F, q, λ) with a fixed fee F and data cap q such that $\lambda r x + y \leq q$.[12] The firm's profits are $\pi = \alpha F - cQ$.

2.2 Usage Decisions

Given a tariff (F, q, λ), we can assume without loss of generality that the cap binds at least weakly (otherwise the ISP could increase its profits by investing less in capacity). Thus subscribers maximize their utility from usage by solving

$$\max_x U(x, q - \lambda r x).$$

the solution $(x^*(\lambda, q), y^*(\lambda, q))$ is given by $y^* = q - \lambda r x^* \geq 0$ (we will concentrate on the case where both services are consumed, i.e. $y^* > 0$) and the first-order condition

$$U_x(x^*, q - \lambda r x^*) = \lambda r U_y(x^*, q - \lambda r x^*)$$

For further reference below, demand for x depends on λ and q as follows:

$$\begin{aligned} x^*_\lambda &= \frac{-r x^* U_{xy} - r U_y + \lambda r^2 x^* U_{yy}}{-\left(U_{xx} - 2\lambda r U_{xy} + \lambda^2 r^2 U_{yy}\right)}, \\ x^*_q &= \frac{U_{xy} - \lambda r U_{yy}}{-\left(U_{xx} - 2\lambda r U_{xy} + \lambda^2 r^2 U_{yy}\right)}. \end{aligned}$$

Note that the strict concavity of $U(.)$ implies that the denominators are strictly positive.[13] On the other hand, neither effect can be signed in general, since the numerators depend on the degrees of substitutability and zero-rating.

[11] For simplity, we allow for intermediate values of λ, as this helps in identifying when zero-rating is profit-maximizing.

[12] Since consumers are homogeneous in usage, tariffs with usage prices either below or above the cap lead to exactly the same outcome.

[13] For any non-zero $a \in \mathbb{R}^2$ we have $a' \left(D^2 U\right) a < 0$; here $a = (1, -\lambda r)$.

The capacity actually used per consumer is

$$k = rx^* + y^* = q + (1 - \lambda) rx^* (\lambda, q),$$

which defines the cap q implicitly as a function of (λ, k). We find

$$q_k = \frac{1}{1 + (1 - \lambda) rx^*_q}, \quad q_\lambda = r\frac{x^* - (1 - \lambda) x^*_\lambda}{1 + (1 - \lambda) rx^*_q}.$$

It would be reasonable to expect $q_k > 0$, i.e. that the cap given to consumers increases with the available capacity per consumer, independently of the degree of zero-rating. This does not follow from the assumptions made so far. Therefore we make the following additional assumption:

$$\text{(M)} \qquad rU_{xy} > U_{xx}.$$

That is, if services are substitutes then they are not too homogeneous.[14] Under assumption (M) it follows that

$$q_k = \frac{-\left(U_{xx} - 2\lambda rU_{xy} + \lambda^2 r^2 U_{yy}\right)}{(1 - \lambda)(rU_{xy} - U_{xx}) - \lambda(U_{xx} - 2rU_{xy} + r^2 U_{yy})} > 0,$$

since the numerator and the second term in the denominator are also positive due to the concavity of U.

2.3 The ISP's Problem

Now we consider the ISP's choices concerning tariff, zero-rating and investments. In the following sections we will analyze the latter two decisions in detail. Let

$$S^* (\lambda, k, \alpha) = S (x^* (\lambda, q (\lambda, k)), y^* (\lambda, q (\lambda, k)), z^* (\lambda, q (\lambda, k))),$$

where $z^* = nx^*$. The ISP solves

$$\max_{\lambda,k,F,\alpha} \pi = \alpha (F - ck) \quad s.t. \ \alpha = G (S^* (\lambda, k, \alpha) - F).$$

The generic outcome of this profit maximization problem is described in the following Lemma. In the next section we analyze the actual choices of zero-rating and capacity investments.

Lemma 1 *At the profit maximum,*[15] *the ISP*

1. *chooses both the degree of zero-rating λ and the per-consumer capacity k such as to maximize per-consumer net surplus including network effects $S^* - ck$;*

[14] In our specific model, this condition becomes $1 - r\gamma > 0$, where $\gamma = -U_{xy} \in (-1, 1)$ indicates the degree of substitutability.

[15] We assume that sufficient second-order conditions for a maximum hold. These do so in our specific model.

2. *and sets the fixed fee as given by ($\varepsilon = FG'/\alpha$ is the subscription elasticity)*

$$\frac{F - (ck - \alpha\beta z_\alpha^*)}{F} = \frac{1}{\varepsilon}.$$

3. *Profits increase with the strength of network effects and decrease with the cost of capacity,*

$$\frac{d\pi}{d\beta} = \alpha z^* > 0, \quad \frac{d\pi}{dc} = -\alpha k < 0.$$

Proof. Consider the Lagrangian

$$L = \alpha (F - ck) + \mu (G (S^* - F) - \alpha),$$

with shadow cost of market share $\mu > 0$. The necessary first-order conditions for an interior maximum are:

$$\begin{aligned} 0 &= \frac{\partial L}{\partial \lambda} = \mu G' S_\lambda^*, \quad 0 = \frac{\partial L}{\partial k} = -\alpha c + \mu G' S_k^*, \\ 0 &= \frac{\partial L}{\partial F} = \alpha - \mu G', \quad 0 = \frac{\partial L}{\partial \alpha} = F - ck + \mu (G' S_\alpha^* - 1). \end{aligned}$$

The above results are obtained from $\mu = \alpha/G'$, which implies $S_\lambda^* = 0$, $S_k^* = c$ and $F = ck - \alpha\beta z_\alpha^* + \alpha/G'$. The boundary maxima at $\lambda = 0$ or $\lambda = 1$ are given by the conditions $S_{\lambda|\lambda=0}^* \leq 0$ and $S_{\lambda|\lambda=1}^* \geq 0$, respectively. Sufficient second-order conditions for a maximum are given by the requirement that the bordered Hessian of the Lagrangian be negative definite. These imply in particular that S^* is strictly concave in (λ, k), i.e. $S_{\lambda\lambda}^*, S_{kk}^* < 0$ and $S_{\lambda\lambda}^* S_{kk}^* - (S_{\lambda k}^*)^2 > 0$, and that network effects are not too strong and subscription demand not too convex, $2 - 2G'\beta n_\alpha x^* - \alpha G''/(G')^2 > 0$.

As for the last statement, the envelope theorem implies that $\frac{d\pi}{d\beta} = \frac{\partial L}{\partial \beta} = \mu G' z^* = \alpha z^*$ and $\frac{d\pi}{dc} = \frac{\partial L}{\partial c} = -\alpha k$. ■

The fact that the ISP charges for subscriptions and not for usage (essentially a two-part tariff with a zero usage price) allows him to decouple the decisions on zero-rating and capacity choice from the actual pricing choices: both are taken to maximize surplus net of capacity costs. This confirms to the usual logic under to two-part tariffs: The subscription fee is used to extract rents, while other decision variables maximize the available surplus per customer.[16]

The expression for fixed fees provides the traditional monopoly pricing formula, with "marginal costs" $ck - \alpha\beta z_\alpha^*$. The latter term translates the network effect benefits from adding a marginal subscriber, which are $\alpha\beta x^* > 0$ with firm-level network effects – and zero with club effects. Thus club effects will lead to a higher subscription price (and lower subscription numbers) than network-level benefits.

The last statement of the proposition shows that the model behaves as one would expect: Network effects benefit the firm, at rate $\alpha^2 x^*$ under firm-level effects and $\alpha v x^*$ under club effects. The effect of capacity cost is the obvious one.

[16] We show below that this is not the same as maximizing total welfare, since it does not take into the account the actual number of subscribers. The latter is chosen such as to maximize profits, not welfare.

3 To Zero-Rate or Not to Zero-Rate

We will now consider the first substantive question of the paper, which is under which conditions the ISP will choose whether to zero-rate the usage of service x. Above we found that profits are maximized at the level of zero-rating that maximizes surplus S^* (we can drop the capacity cost here).

The first point to take note of is that the purpose of zero-rating in this context is to increase the consumption of service x, i.e. change users' consumption pattern. A direct effect of changing the consumption pattern is a reduction in consumer surplus as compared to users' (non-zero-rated) usage given the same capacity per user, since users would have chosen the latter and not the former usage allocation. Thus zero-rating only makes sense for the ISP if it brings benefits from other sources, such as network effects (or other externalities that users do not take into account) or payments from content providers or advertisers.

Thus the choice of zero-rating involves a trade-off between a consumption distortion and the benefits resulting from a higher consumption of some service(s). We find exactly this, as described in the following Proposition:

Proposition 1 *(Zero-rating)*

1. *The ISP (fully) zero-rates service x if and only if network effects are strong enough,*

$$\beta n \geq rU_y. \tag{1}$$

2. *A joint data cap, rather than partial zero-rating, is optimal if and only if there are no network effects,*

$$\beta = 0.$$

Proof. The derivative of surplus S^* with respect to the degree of zero-rating is

$$\begin{aligned} S^*_\lambda &= \frac{d}{d\lambda} S\left(x^*\left(\lambda, q\left(\lambda, k\right)\right), y^*\left(\lambda, q\left(\lambda, k\right)\right), nx^*\left(\lambda, q\left(\lambda, k\right)\right)\right) \\ &= \left(U_x + \beta n\right)\frac{dx^*}{d\lambda} + U_y \frac{dy^*}{d\lambda}. \end{aligned}$$

The condition $rx^* + y^* = k$ implies that $\frac{dy^*}{d\lambda} = -r\frac{dx^*}{d\lambda}$, which together with $\frac{dx^*}{d\lambda} = x^*_\lambda + x^*_q q_\lambda$ leads us to

$$S^*_\lambda = \frac{\left(U_x + \beta n - rU_y\right)\left(x^*_\lambda + rx^* x^*_q\right)}{1 + (1-\lambda) r x^*_q} = \frac{\left(U_x + \beta n - rU_y\right) rU_y q_k}{U_{xx} - 2\lambda r U_{xy} + \lambda^2 r^2 U_{yy}}.$$

The denominator is negative due to the strict concavity of U, and $rU_y q_k > 0$ due to assumption (M) and the binding cap. With full zero-rating, $\lambda = 0$, users consume service x at their bliss-point $\bar{x}(y^*)$, thus $U_x = 0$. Since $\lambda = 0$ is optimal if $S^*_{\lambda|\lambda=0} \leq 0$, this implies that we must have $\beta n \geq rU_y$.

On the other hand, at $\lambda = 1$ (joint cap) optimal usage implies $U_x = rU_y$, thus unless $\beta = 0$ we have (at a binding cap) $S^*_{\lambda|\lambda=1} < 0$, i.e. the maximum is found at some $\lambda < 1$. ■

Condition (1) provides a simple summary of the relevant trade-offs between consumption of the two services. On the one hand, one more unit of good x provides zero direct consumption benefits at the bliss point, but provides network effects as measured by $n\beta$. Remember that with $n = \alpha$ these effects are proportional to the user base, while with $n = v$ they are proportional to club size. On the other hand, the costs of zero-rating are given by the size of the consumption distortion as measured by rU_y, the opportunity cost of not consuming r additional units of y. This opportunity cost decreases with the available capacity per customer:

$$\left.\frac{dU_y}{dk}\right|_{\lambda=0} = \left(U_{xy}x^*_q + U_{yy}y^*_q\right) q_k\big|_{\lambda=0} = -\frac{U_{xx}U_{yy} - U^2_{xy}}{rU_{xy} - U_{xx}},$$

which is negative due to the concavity of U and assumption (M). Thus the ISP is more likely to adopt zero-rating if his capacity per consumer is high. Equally, the ISP is more likely to zero-rate a certain service if its capacity usage r is low, as was found by Dotecon's survey of European zero-rating offers (EC 2017, p.16).

The result for the joint cap further illustrates the above discussion. Clearly, if there are no benefits from network effects (or other externalities or payments from third parties), the ISP has no reason to adopt zero-rating: Surplus and profits are maximized if consumers choose by themselves how of each service to consume below the cap.

4 Capacity Investment

We will now consider investment in capacity, for any $\lambda \in [0, 1]$. Above we showed the latter is chosen to maximize net surplus $S^* - ck$, which together with the optimality conditions $U_x = \lambda rU_y$ and $y^*_q = 1 - \lambda rx^*_q$ yields the first-order condition

$$\begin{aligned} c &= S^*_k = \frac{d}{dk}S\left(x^*\left(\lambda, q\left(\lambda, k\right)\right), y^*\left(\lambda, q\left(\lambda, k\right)\right), nx^*\left(\lambda, q\left(\lambda, k\right)\right)\right) \\ &= \left[\left(U_x + \beta n\right) x^*_q + U_y y^*_q\right] q_k = \left[\beta n x^*_q + U_y\right] q_k \end{aligned}$$

Again, the trade-off becomes visible: The optimal capacity is achieved when the marginal costs of capacity are equal to the marginal benefits from network effects plus higher consumption of the non-zero-rated service y. Actually, the network effects term is not necessarily conducive to higher investment, on the contrary: Assuming an interior maximum with $S^*_{kk} < 0$, the effect of network effects on the optimal investment is given by

$$k_\beta = -\frac{S^*_{k\beta}}{S^*_{kk}} = \frac{nx^*_q q_k}{-S^*_{kk}} = \frac{\left(U_{xy} - \lambda rU_{yy}\right) nq_k}{S^*_{kk}\left(U_{xx} - 2\lambda rU_{xy} + \lambda^2 r^2 U_{yy}\right)},$$

the sign of which is equal to that of the expression $(U_{xy} - \lambda rU_{yy})$. That is, stronger network effects increase capacity investment if and only if $U_{xy} > \lambda rU_{yy}$, i.e. services are complements or at least not very strong substitutes. For zero-rating, i.e. $\lambda = 0$, we immediately obtain the following result:

Proposition 2 *Under zero-rating, stronger network effects increase capacity investment if services are complements, and decrease it if they are substitutes.*

Proof. For $\lambda = 0$, the sign of k_β is equal to that of U_{xy}. ■

It seems natural that stronger network effects should lead to higher capacity investments. Under complements, the ISP will indeed want to encourage higher consumption of all services, including the non-zero-rated one, and therefore increases capacity. With substitutes, however, the ISP increases users' bliss points for service x by depressing their consumption of service y through a lower data cap. As a result, total data consumption is lower and the ISPs invests less in capacity.

Under a joint cap, stronger network effects are more likely to increase capacity investment ($U_{yy} < 0$), since there the ISP does not restrict users' consumption of the other service.

A second interesting issue is whether zero-rating or a joint cap lead to higher investment. While only one of the two choices is optimal for the ISP, this may be a relevant question when a prohibition of zero-rating is considered. It turns out that no unambiguous answer can be given in general. Therefore we now use the specific model set out in the appendix to provide some intuitions.

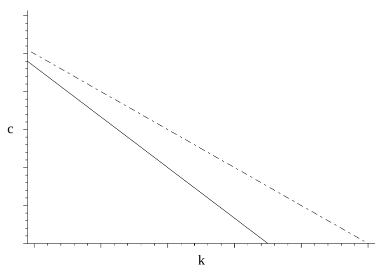

Figure 1: $\beta n > r\,(1-\gamma)$

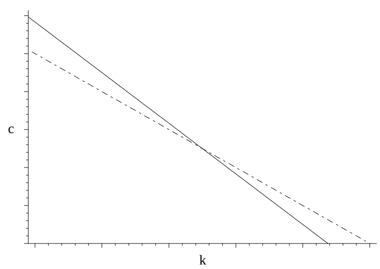

Figure 2: $\beta n < r\,(1-\gamma)$

Figures 1 and 2 depict S_k^* for zero-rating (continuous line) and a joint cap (dotted line), which cross at some $\bar{k}$ such that $\beta n = rU_y$. For strong network effects (Figure 1) the ISP chooses zero-rating, which then implies a lower capacity investment than a joint cap. For weaker network effects (Figure 2) the outcome depends on the investment cost: For c below the level of the intersection, the ISP chooses zero-rating; for levels of c above it, zero-rating is not optimal (some partial degree of zero-rating would be profit-maximizing). Again the capacity investment chosen is the lower one. This implies that due to a better exploitation of network effects profits and welfare are higher even at a lower level of capacity.

5 Welfare and Consumer Surplus

Total welfare in our model is obtained from summing profits $\pi = \alpha(F - ck)$ and consumer surplus $CS = \alpha(S^* - F)$, resulting in $W = \alpha(S^* - ck)$. Welfare depends on both the surplus per subscriber and the number of subscribers. This does not take into account that there may be further externalities, on content providers, advertisers, sellers of goods and services, or society at large. To capture these, define

$$\tilde{S}^* = S^* + E + \eta_x x^* + \eta_y y^*,$$

where E is a subscription externality, and η_x, η_y measure the strength of consumption externalities on third parties for services x and y, respectively. Welfare then becomes $W = \alpha\left(\tilde{S}^* - ck\right)$. In order to separate these issues, we will first consider the case without such externalities.

Remember that the ISP maximizes per-subscriber surplus and chooses the subscriber number separately in order to maximize profits. Thus it does not explicitly take into account the social welfare effect in this latter choice, and these choices may not be socially optimal.

We find, in the case without further externalities ($E = \eta_x = \eta_y = 0$):

Proposition 3 *Social optimality of the ISP's choices of zero-rating and capacity in the absence of externalities:*

1. *The ISP's choices of zero-rating and capacity investment are socially optimal if either there are no network effects ($\beta = 0$), or these are club effects ($n = v$);*

2. *When the ISP chooses zero-rating under firm-level network effects ($n = \alpha$), it is socially optimal. Capacity investment is socially optimal if either services are independent ($U_{xy} = 0$), or if the market is fully covered and services are complements ($U_{xy} > 0$). Otherwise the ISP underinvests (with complements) or overinvests (with substitutes).*

3. *With firm-level network effects and a joint cap, the ISP underinvests for a larger range of the parameter space.*

Proof. Relegated to Appendix B. ■

The first observation is that there are cases where the ISP's decisions concerning zero-rating and capacity investment do not influence the profit-maximizing number of subscribers. This happens when either there are no network effects or if there are club network effects. In both cases the externality between consumers does not depend on the ISP's decisions, and thus there will be no wedge between profit-maximizing and socially optimal choices.

Under firm-level network effects this is different, and here the potential for suboptimal decisions arises. Still, if the ISP chooses zero-rating, then a social planner would do the same. The reason is that "more" zero-rating (lower λ) increases the consumption of the zero-rated good, which raises social surplus overall since network effects are strong enough.

The situation is more complicated concerning capacity investments. If the market is not fully covered then the ISP's capacity choice is only socially optimal if the services are independent in demand. If they are substitutes then it would be socially optimal to move more consumption to the zero-rated good, reducing capacity while doing so; with complements the opposite is true.

Finally, with firm-level network effects and a joint cap it is socially optimal to invest more in capacity, thus there are more cases when underinvestment occurs. While a joint cap would not be profit-maximizing if $\beta > 0$, this situation will arise if zero-rating was prohibited.

If we now consider the effect of including subscription and usage externalities, we obtain the following conclusions:

Proposition 4 *Social optimality of the ISP's choices of zero-rating and capacity with externalities:*

1. *Positive subscription externalities under firm-level network effects ($n_\alpha E > 0$), and sufficiently strong usage externalities on service x ($\eta_x > r\eta_y$) strengthen the social optimality of zero-rating, while they have no effect on the social optimality of the investment decision.*

2. *A large usage externality on the non-zero-rated service y may (but need not) imply that zero-rating is not optimal or that capacity investment is too low.*

Proof. In the proof of Proposition 3 it is shown that with $E = \eta_x = \eta_y = 0$ we obtain $\frac{dW}{dh} = \beta n_\alpha \Phi \left(S^* - ck\right) \frac{dx^*}{dh}$ for $h = k, \lambda$, where $n_\alpha \in \{0, 1\}$ and $\Phi > 0$. Taking into account externalities, we obtain

$$\frac{dW}{dh} = \frac{d\alpha}{dk}\left(\tilde{S}^* - ck\right) + \alpha\left(S_k^* - c + \eta_x \frac{dx^*}{dk} + \eta_y \frac{dy^*}{dk}\right),$$

which can be written as

$$\begin{aligned} \frac{dW}{d\lambda} &= \left(\beta n_\alpha \Phi\left(\tilde{S}^* - ck\right) + \alpha\left(\eta_x - r\eta_y\right)\right)\frac{dx^*}{d\lambda}, \\ \frac{dW}{dk} &= \left(\beta n_\alpha \Phi\left(\tilde{S}^* - ck\right) + \alpha\left(\eta_x - r\eta_y\right)\right)\frac{dx^*}{dk} + \alpha\eta_y. \end{aligned}$$

Both $n_\alpha E > 0$ and $\eta_x > r\eta_y$ then strengthen the social optimality of $\lambda = 0$, but do not affect the optimality of the investment decision (since the sign of the term in parentheses does not become negative). On the other hand, a large $\eta_y > 0$ may imply that zero-rating is not optimal (if $\beta n_\alpha \Phi y^* < \alpha r$), and underinvestment in capacity. ■

The additional terms in the expression for social welfare on the one hand increase welfare per subscriber and therefore raise the socially optimal number of subscribers. This favours zero-rating. On the other hand, the effect of usage externalities depend on whether the externalities are stronger with respect to the zero-rated or non-zero-rated service, as one would expect.

A related issue is whether consumers benefit or not from an ISP's decision to adopt zero-rating due to network effects. While zero-rating increases per-consumer surplus if network effects are strong, this surplus is at least partially extracted through subscription payments, and, as mentioned above, zero-rating itself creates a distortion in the relative consumption of the two services. Therefore it is not immediately obvious that zero-rating should increase consumer surplus; but the following Proposition shows that this is indeed the case:

Proposition 5 *If the ISP adopts zero-rating due to network effects, consumer surplus either remains constant (if either $\beta = 0$ or with club effects) or increases (under firm-level network effects).*

Proof. Consumer surplus is $CS = \alpha (S^* - F)$. We have

$$\begin{aligned} \frac{dCS}{d\lambda} &= (S^* - F) G' \frac{d(S^* - F)}{d\lambda} + \frac{d(S^* - F)}{d\lambda} G(S^* - F) \\ &= \frac{d\alpha}{d\lambda} \left(S^* - F + \frac{\alpha}{G'} \right) = \beta n_\alpha \frac{dx^*}{d\lambda} \Phi (S^* - ck + \alpha\beta n_\alpha x^*) . \end{aligned}$$

Since the term in parentheses is positive and $dx^*/d\lambda < 0$, either $\beta n_\alpha = 0$ or $dCS/d\lambda < 0$ at $\lambda = 0$. ■

Essentially, any action that increases net surplus $S^* - F$ per consumer also increases participation and therefore has a double effect on consumer surplus. One such action is the choice of zero-rating – but only if network effects are at the firm level, because then the ISP has an additional incentive to increase subscriber numbers. With club effects the ISP simply sets a higher fixed fee which captures the rents from network effects, but consumers do not gain from this.

The general take-away from this is that if ISPs adopt zero-rating based on club or firm-level network effects this increases welfare, while the optimality of the capacity investment depends on demand-side features. With club effects all the gains go to the ISP, while with firm-level network effects also consumer surplus increases. Including further externalities does not fundamentally change these conclusions.

6 Zero-Rating and Competition

Now we assume that two ISPs i and j compete, offering tariffs (F_i, q_i, λ_i), and that consumers' utility of usage of services x and y is

$$S_i (x_i, y_i; z_i, z_j) = U_i (x_i, y_i) + \beta_i z_i + \zeta_i z_j,$$

where $\beta_i \geq \zeta_i \geq 0$ and $z_l = n_l x_l$, $l = i, k$, with $n_l = \alpha_l$ or v. As above consumers maximize their utility subject to the cap, $\lambda_i r x_i + y_i \leq q_i$, resulting in gross surplus $S_i^* (\lambda_i, k_i, \alpha_i, z_j) = U_i (x_i^*, y_i^*) + \beta_i n_i x_i^* + \zeta_i z_j^*$ and net surplus $w_i = S_i^* - F_i$.

The subscription demand of ISP i is given by $\alpha_i = G_i (w_i, w_j)$, with partial derivatives $G_{iw_i} > -G_{iw_j} > 0$, i.e. subscriptions are substitutes and the own-price effect is stronger than the cross-price effect. Let $\sigma_i = G_{iw_j}/G_{iw_i} \in (-1, 0)$ denote the displacement ratio. We consider the Nash equilibrium where ISP simultaneously maximize their profits over all variables, given (λ_j, k_j, F_j):

$$\max_{\lambda_i, k_i, F_i, \alpha_i} \pi_i = \alpha_i \left(F_i - c_i k_i\right) \quad s.t.\ \alpha_i = G_i \left(S_i^* - F_i, S_j^* - F_j\right).$$

Following the same steps as in Lemma 1, we arrive at the following results:

Lemma 2 *In the duopoly Nash equilibrium, ISP i chooses the degree of zero-rating λ_i and per-consumer capacity k_i such as to maximize $S_i^* - c_i k + \sigma_i S_j^*$.*

In choosing zero-rating or capacity, each ISP does not only take into account their effects on the surplus provided to its own customers, but also how much subscription demand is displaced to the other ISP through spillovers of network effects. Following the same steps as in Proposition 1, we prove the following:

Proposition 6 *Let $\tilde{\beta}_i = \left(\beta_i + \sigma_i \zeta_j\right)$. In the Nash equilibrium, ISP i will choose zero-rating if*

$$\tilde{\beta}_i n_i \geq r U_{iy_i},$$

which implies in particular that

1. *Zero-rating is less likely to arise ($\tilde{\beta}_i$ is lower) if subscriptions are close substitutes ($\sigma_i \approx -1$) and network effect spillovers are strong (ζ_j large);*
2. *With firm-level network effects ($n_i = \alpha_i$), larger networks are more likely to adopt zero-rating.*

This decision is mediated by $\tilde{\beta}_i = \left(\beta_i + \sigma_i \zeta_j\right)$, the "residual network effect", as are the choices of fixed fees ($\varepsilon_i = F_i G_{iw_i}/\alpha_i$),

$$\frac{F_i - \left(c_i k - \alpha_i \tilde{\beta}_i n_{i\alpha_i} x_i^*\right)}{F_i} = \frac{1}{\varepsilon_i}$$

and the optimal capacity:

$$c_i = S_{ik_i}^* + \sigma_i S_{jk_i}^* = \left[\tilde{\beta}_i n_i x_{iq_i}^* + U_{iy_i}\right] q_{ik_i}.$$

Similar to the case of monopoly, under zero-rating capacity investment increases in $\tilde{\beta}_i$ if services x and y are complements, and decreases it if they are substitutes:

$$k_{\tilde{\beta}_i} = \frac{1}{-\left(S_{ik_i}^* + \sigma_i S_{jk_i}^*\right)_{k_i}} \frac{n_i U_{ix_i y_i}}{r U_{ix_i y_i} - U_{ix_i x_i}}.$$

Thus the results derived with a monopoly ISP carry over to the case of duopoly, with the only change being that instead of the "gross network effect" β_i now we need to consider the residual network effect $\tilde{\beta}_i$. The latter captures the trade-off between three different forces: firm-level network effects, spillovers, and brand differentiation.

7 Conclusions

In this paper we have explored an alternative explanation for the rise of zero-rating tariffs which does not hinge on payments from content providers to internet service providers (ISPs). Rather, ISPs can use zero-rating to better exploit network effects on certain services, which happens to maximize their profits from subscriptions if these network effects are strong enough. If zero-rating is chosen it will also be socially optimal, unless, possibly, with a very strong externality on third parties of the non-zero-rated service.

The optimality condition for zero-rating carries over to duopoly, with the qualification that in this case either network effect spillovers need to be weak or brand differentiation strong enough, in order not to outweigh the firm-level network effects. These three forces are captured succinctly in a measure of the "residual network effect".

Capacity investment is lower under zero-rating than under a joint cap on usage if investment cost is low enough so that zero-rating is adopted. Whether it is at the socially optimal level depends on the type of network effects and the level of market coverage. Stronger network effects may raise or lower capacity investment: This depends on the degree of substitutability between the services offered at each ISP. Under zero-rating, the correspondence is exact: with substitutes (complements) stronger network effects lower (raise) capacity investment.

Further research will consider various issues. For once, we have assumed contents, including their differentiation, capacity usage and network effects, as given. All these factors can be analyzed as choice variables of content providers. These will result in an equilibrium mix of content types which interacts with ISPs' decisions to zero-rate certain services. Second, our setting can be combined with previous models in order to include payments from subscribers to content providers, payments from content providers to ISPs, and issues of market power, into the discussion.

References

[1] BEREC (2016). BEREC Guidelines on the Implementation by National Regulators of European Net Neutrality Rules - BoR (16) 127, http://berec.europa.eu/eng/document_register/subject_matter/berec/download/0/6160-berec-guidelines-on-the-implementation-b_0.pdf.

[2] BEREC (2017). Report on the implementation of Regulation (EU) 2015/2120 and BEREC Net Neutrality Guidelines - BoR (17) 240, http://berec.europa.eu/eng/document_register/subject_matter/berec/download/0/7529-berec-report-on-the-implementation-of-re_0.pdf, December.

[3] EC (2017). Zero-rating practices in Broadband markets, Report by Dotecon. http://ec.europa.eu/competition/publications/reports/kd0217687enn.pdf, February.

[4] Eisenach, J.A. (2015). The Economics of Zero Rating, NERA Economic Consulting, http://www.nera.com/content/dam/nera/publications/2015/EconomicsofZeroRating.pdf, March.

[5] EU (2015). Regulation (EU) 2015/2120 of the European Parliament and the of the Council of 25 November 2015, laying down measures concerning open internet access and amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services and Regulation (EU) No 531/2012 on roaming on public mobile communications networks within the Union ("Open internet regulation"), http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32015R2120.

[6] FCC (2015). "2015 Open Internet Order", https://apps.fcc.gov/edocs_public/attachmatch/FCC-15-24A1.pdf.

[7] FCC (2017). Policy Review of Mobile Broadband Operators' Sponsored Data Offerings for Zero-Rated Content and Services, https://apps.fcc.gov/edocs_public/attachmatch/DOC-342987A1.pdf, January.

[8] FCC (2018). "Restoring Internet Freedom Order", https://www.fcc.gov/document/fcc-releases-restoring-internet-freedom-order.

[9] Greenstein, S., M. Peitz, and T. Valletti (2016). Net neutrality: A fast lane to understanding the trade-offs, The Journal of Economic Perspectives, 30(2), 127–149.

[10] Hoernig, S. (2007). On-net and off-net pricing on asymmetric telecommunications networks, Information Economics and Policy 19 (2007) 171–188.

[11] Hoernig, S. (2014). Competition between multiple asymmetric networks: Theory and applications, International Journal of Industrial Organization, 32, 57–69.

[12] Jaunaux, L., & M. Lebourges (2018). Zero rating and end-users' freedom of choice: an economic analysis, mimeo Orange, Regulatory Department.

[13] Jeon, D.S., Laffont, J.J. and Tirole, J. (2004). On the 'receiver-pays' principle, Rand Journal of Economics, vol. 35(1), pp. 85–110.

[14] Jullien, B. and W. Sand-Zantman (2017). "Internet Regulation, Two-Sided Pricing, and Sponsored Data", TSE Working Paper, n. 12-327, https://www.tse-fr.eu/publications/internet-regulation-two-sided-pricing-and-sponsored-data, December.

[15] Kramer, J. and M. Peitz (2018). A Fresh Look at Zero-Rating, CERRE report, March, http://www.cerre.eu/sites/cerre/files/CERRE_ZeroRating_FinalReport.pdf.

[16] Ofcom (2017). Monitoring compliance with the EU Net Neutrality regulation: A report to the European Commission, https://www.ofcom.org.uk/__data/assets/pdf_file/0018/103257/ net-neutrality.pdf, December.

[17] Schmutzler, A. (2013). Competition and investment – A unified approach, International Journal of Industrial Organization, 31, 477–487.

[18] Schnurr, D. & L. Wiewiorra (2018). Bit-by-Bit Towards Unlimited: An Analysis of Zero Rating and Sponsored Data Practices of Internet Service Providers, mimeo.

[19] Somogyi, R. (2017). The Economics of Zero-rating and Net Neutrality. https://sites.google.com/site/robertsomogyi/Zero_rating_Somogyi.pdf, mimeo, November.

[20] Yoo, C. S. (2017). Avoiding the Pitfalls of Net Uniformity: Zero Rating and Nondiscrimination, Review of Industrial Organization, 50, 509–536.

Appendix A: The Specific Model

Here we set out a specific model of consumption utility and derive its implications. While consumers are heterogeneous in subscription benefits, they all obtain the same utility from actually consuming services, with utility function

$$U(x,y) = x + y - \frac{1}{2}x^2 - \frac{1}{2}y^2 - \gamma xy.$$

The parameter $\gamma \in (-1,1)$ describes the degree of product differentiation, with $\gamma > 0$ for substitutes, $\gamma = 0$ for independent services, and $\gamma < 0$ for complements. In terms of the general notation used above, we have

$$\begin{aligned} U_x &= 1 - x - \gamma y,\ U_y = 1 - y - \gamma x, \\ U_{xx} &= -1,\ U_{xy} = -\gamma,\ U_{yy} = -1. \end{aligned}$$

Consumers maximize this utility subject to the cap $\lambda r x + y \leq q$, and assuming that the cap is binding we obtain the following optimal consumption:

$$x^* = \frac{1 - \lambda r + (\lambda r - \gamma) q}{1 - 2\lambda r\gamma + \lambda^2 r^2},\ y^* = q - \lambda r x^*.$$

From the total data usage $rx^* + y^* = k$ we can solve for the corresponding cap

$$q = \frac{1 - 2\lambda r\gamma + \lambda^2 r^2}{1 - r\gamma + \lambda r (r - \gamma)} k + \frac{r(1-\lambda)(\lambda r - 1)}{1 - r\gamma + \lambda r (r - \gamma)}.$$

Assumption (M) becomes $1 - r\gamma > 0$, which we have seen to be sufficient for $q_k > 0$. Let $K = \frac{1+r}{1+\gamma}$, which is the maximal amount of data that a consumer will use if the cap is not binding. Thus we only need to consider $k \leq K$. On the other hand, $y^* \geq 0$ implies $k \geq r\frac{1-\lambda r}{1-\lambda r\gamma}$ (which is decreasing in λ), i.e. if the capacity per consumer is too small then the non-zero-rated service is crowded out. In the following we concentrate on the case where capacity is above this limit.

Total surplus per consumer at the optimal consumption level can now be written as

$$S^* = \frac{1}{1+\gamma} - \frac{(1-\gamma^2)\left(1 - 2\lambda r\gamma + \lambda^2 r^2\right)}{2\left(1 - r\gamma + \lambda r (r - \gamma)\right)^2}(K-k)^2 + \beta n \frac{1 - \gamma k - \lambda r (1-k)}{1 - r\gamma + \lambda r (r - \gamma)}$$

Maximizing S^* with respect to λ leads to the interior solution

$$\lambda^* = \frac{r\left(1-\gamma^2\right)(K-k) - \beta n\left(1-r\gamma\right)}{r\left(1-\gamma^2\right)(K-k) + \beta nr\left(r-\gamma\right)}.$$

A joint cap ($\lambda = 1$) is optimal only for $\beta = 0$, while zero-rating ($\lambda = 0$) is optimal if $\beta n \geq \frac{r\left(1-\gamma^2\right)}{1-r\gamma}(K-k)$. Note that the latter threshold is weaker if capacity per consumer is larger, i.e. cheaper capacity investment increases the chance that zero-rating is optimal.

As for optimal investment, the condition $S_k^* = c$ becomes

$$c = \frac{\left(1-2\lambda r\gamma + \lambda^2 r^2\right)\left(1-\gamma^2\right)}{\left(1-r\gamma+\lambda r\left(r-\gamma\right)\right)^2}(K-k) + \beta n\frac{\lambda r-\gamma}{1-r\gamma+\lambda r\left(r-\gamma\right)}.$$

The right-hand side is strictly decreasing in k, and thus $S_{kk}^* < 0$ and there is at most one solution. Network effects shift investment up if services are either complements or sufficiently weak substitutes - otherwise they shift it down.

Now we compare the conditions $S_k^* = c$ for $\lambda = 0$ and $\lambda = 1$. We obtain for zero-rating

$$c = C^{ZR}(k) = \frac{1-\gamma^2}{1-2r\gamma+r^2\gamma^2}(K-k) + \beta n\frac{-\gamma}{1-r\gamma},$$

defined on $r \leq k \leq K$, and for a joint cap

$$c = C^{JC}(k) = \frac{1-\gamma^2}{1-2r\gamma+r^2}(K-k) + \beta n\frac{r-\gamma}{1-2r\gamma+r^2},$$

defined on $r\frac{1-r}{1-r\gamma} \leq k \leq K$, where the left boundary is smaller than r. Thus we consider the interval $k \in [r, K]$. The curve C^{ZR} has a more negative slope, which means that for any point $\bar{k}$ such that $C^{ZR}\left(\bar{k}\right) = C^{JC}\left(\bar{k}\right)$ then $C^{ZR}(k) < C^{JC}(k)$ for all $k > \bar{k}$, i.e. given a marginal cost of capacity c investment would be higher under joint caps in this range.

Both curves are represented in Figures 1 and 2 in the text. Remember that the threshold on network effects depends negatively on the available capacity k. Starting with the lower boundary on k, if $\beta n > \frac{r\left(1-\gamma^2\right)}{1-r\gamma}(K-r) = r(1-\gamma)$ then $C^{ZR}(k) < C^{JC}(k)$ for all $k \in [r, K]$, that is, capacity investment would be higher under a joint cap, but the firm chooses zero-rating and the correspondingly lower capacity investment through $c = C^{ZR}(k)$. For lower values of βn there is a $\bar{k} \in [r, K]$ with $C^{ZR}\left(\bar{k}\right) = C^{JC}\left(\bar{k}\right)$ exactly such that $\beta n = \frac{r\left(1-\gamma^2\right)}{1-r\gamma}\left(K-\bar{k}\right)$. Denote $\bar{c} = C^{ZR}\left(\bar{k}\right)$ the level of cost for which $\bar{k}$ would be the optimal investment. Then for $c < \bar{c}$ the ISP chooses zero-rating and capacity according to $c = C^{ZR}(k) < C^{JC}(k)$, and for $c > \bar{c}$ zero-rating is not chosen (Here partial zero-rating is optimal, as long as $\beta > 0$).

Appendix B: Longer Proofs

Proof of Proposition 3:

Proof. We will consider both zero-rating and capacity simultaneously by letting $h \in \{\lambda, k\}$ and considering dW/dh at the ISP's optimal choice given by $d\left(S^* - ck\right)/dh = 0$, with

$$\frac{dW}{dh} = \frac{d\alpha}{dh}\left(S^* - ck\right) + \alpha\frac{d\left(S^* - ck\right)}{dh} = \frac{d\alpha}{dh}\left(S^* - ck\right).$$

Thus if per-consumer surplus is positive then the social optimality of the chosen level of h depends on the sign of $d\alpha/dh$. With $\alpha = G(S^* - F)$ and $F = ck - \alpha\beta n_\alpha x^* + \alpha/G'$ we have

$$\begin{aligned}
\frac{d\alpha}{dh} &= \frac{d(S^* - F)}{dh} G' \\
&= \left(\frac{d(S^* - ck)}{dh} + \alpha\beta n_\alpha \frac{dx^*}{dh} + \frac{d\alpha}{dh}\beta n_\alpha x^* - \frac{\frac{d\alpha}{dh}}{G'} + \frac{\alpha G'' \frac{d(S^*-F)}{dh}}{(G')^2} \right) G' \\
&= \alpha\beta n_\alpha \frac{dx^*}{dh} G' - \frac{d\alpha}{dh}\left(1 - \beta n_\alpha x^* G' - \frac{\alpha G''}{(G')^2} \right),
\end{aligned}$$

which results in

$$\frac{d\alpha}{dh} = \beta n_\alpha \frac{dx^*}{dh}\Phi, \text{ with } \Phi \equiv \frac{\alpha G'}{2 - \beta n_\alpha x^* G' - \alpha G''/(G')^2} > 0.$$

Here $\Phi > 0$ is implied by the sufficient second-order conditions mentioned above. Clearly $d\alpha/dh = 0$ (the ISP's decisions are socially optimal) if either there are no network effects ($\beta = 0$) or if these are given by club effects ($\beta > 0$ but $n_\alpha = 0$).

Consider now firm-level network effects, i.e. $\beta > 0$ and $n_\alpha = 1$. Then dW/dh has the signs of

$$\begin{aligned}
\frac{dx^*}{d\lambda} &= \frac{rU_y q_k}{U_{xx} - 2\lambda r U_{xy} + \lambda^2 r^2 U_{yy}} < 0, \\
\frac{dx^*}{dk} &= \frac{U_{xy} - \lambda r U_{yy}}{(1-\lambda)(rU_{xy} - U_{xx}) - \lambda(U_{xx} - 2rU_{xy} + r^2 U_{yy})},
\end{aligned}$$

where the latter has the sign of $U_{xy} - \lambda r U_{yy}$.

Since $dx^*/d\lambda < 0$ we have $dW/d\lambda < 0$ at the ISP's optimal choice, which implies that it is socially optimal to have "more" zero-rating than the ISP adopts – in particular, a profit-maximizing choice of $\lambda = 0$ is then also socially optimal. This effect arises because network effects increase subscriber numbers.

Now we consider capacity choice. If the ISP chooses zero-rating then dW/dk at his optimal choice has the sign of U_{xy}, i.e. is positive with complements and negative with substitutes. That is, the ISP underinvests or overinvests in capacity, respectively, unless services are independent or the whole market is covered (in which case higher investment cannot increase customer numbers). With a joint cap, on the other hand, dW/dk has the sign of $U_{xy} - rU_{yy}$, which positive even with weak substitutes. Therefore the ISP underinvests for a larger range of the parameter space. ■

INOVA

NOVA School of Business & Economics

Shaping powerful minds

Nova School of Business and Economics
Faculdade de Economia
Universidade Nova de Lisboa
Campus de Campolide
1099-032 Lisboa PORTUGAL
Tel.: +351 213 801 600

www.novasbe.pt