Zero-Rating and Network Effects*

Steffen Hoernig
Nova School of Business and Economics†

Francisco Monteiro
Nova School of Business and Economics‡

October 30, 2019

Abstract

We consider internet service providers' incentives to zero-rate, i.e. to not count the usage of certain services towards data allowances, in the absence of payments from content providers. We show that zero-rating is adopted if and only if it strongly increases subscriptions. For this it is necessary that participation (as opposed to usage) network effects are strong enough and if zero-rating offers raise expectations about other subscribers' usage. Zero-rating then also maximizes total welfare.

JEL Classification: D21; L51; L96.

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

---

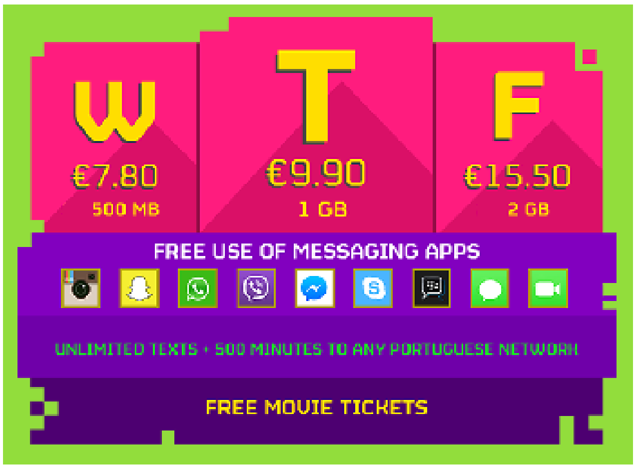
*This is a simultaneously more general and condensed version of a previous paper circulated as "Zero-Rating, Network Effects and Investment Incentives". We would like to thank Marc Bourreau and audiences at PEJ 2018, EARIE 2018, and FSR Competition and Media 2019.

†Campus de Carcavelos, Rua da Holanda, n.1, 2775-405 Carcavelos, Portugal, shoernig@novasbe.pt.
Steffen Hoernig was funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, LISBOA-01-0145-FEDER-022209), POR Lisboa (LISBOA-01-0145-FEDER-007722, LISBOA-01-0145-FEDER-022209) and POR Norte (LISBOA-01-0145-FEDER-022209).

‡This paper arose from Francisco' MSc Thesis at Nova SBE.

1 Introduction

Mobile broadband subscriptions tend to come with a cap on monthly data usage. The term "zero-rating" refers to the recent practice of mobile internet service providers (ISPs) to exempt traffic for certain services from these caps, such as social network and messaging applications (Facebook, WhatsApp, Instagram, Twitter), and video services (Netflix).



Offer of June 2018 by Portuguese mobile operator NOS for the youth segment, zero-rating messaging apps.

However, the practice of zero-rating is controversial, as it may conflict with the principle of "net neutrality", which states that all internet traffic, in particular, traffic by companies providing services via the internet, such as news and search sites, social networks, messaging and data management, in short "content providers" (CPs), should be treated equally.[1] While the U.S. FCC in 2015 allowed for a case-by-case review of zero-rating due to potential consumer benefits, the Netherlands in 2016 forbid pricing schemes such as zero-rating.

Zero-rating has been the subject of some recent attention in the academic literature. 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. Jaunaux and Lebourges (2019) 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. Other papers deal explicitly with payments from CPs to ISPs as the motive for zero-rating, such as Jullien and Sand-Zantman (2018) and Somogyi (2017), while Schnurr and Wiewiorra (2018) and Gautier and Somogyi (2018) consider it as a price discrimination device.

In Europe payments from CPs to ISPs are illegal under the present net neutrality regulations, but still we see zero-rating offers proliferating, in particular for messaging and social network services. Thus other factors must make zero-rating attractive to ISPs.

[1] See, e.g., Greenstein, Peitz and Valletti (2016).

Therefore, and contrary to the previous literature, we exclude payments from content providers to ISPs as the reason for the zero-rating of certain services. Instead, we follow up on the observation that social networking and messaging apps are strongly represented among the zero-rated services. These give rise to network externalities, i.e. each single user's consumption increases other users' surplus. Zero-rating can be used to increase usage and thus total surplus – which makes the ISP's offer more appealing and raises participation.[2]

We show that network effects can be a sufficiently strong motive for zero-rating, and that both consumers and society benefit from it.[3] Zero-rating is a means to raise consumer *expectations about other subscribers' usage*, which in turn increases their willingness to subscribe in the first place. It therefore serves as a coordination device to increase aggregate usage and participation, allocating sparse network capacity efficiently at the same time. Without this effect on expectations and participation, zero-rating only distorts consumption patterns, lowering surplus and profits.

We provide a model with network effects for a monopoly ISP with general specifications of subscription and usage demands, and capacity constraints. The ISP decides whether to adopt zero-rating or not. Zero-rating is modeled 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 conditions for the trade-offs involved. In doing so, we distinguish between *participation network effects*, which is the increase in the benefits from subscribing, and *usage network effects*, which is the increase in the benefits of more usage – only the former counts.

Our main results are as follows: Zero-rating is indeed a profit-maximizing choice for the ISP if the participation network effect is strong enough, and if potential subscribers understand that zero-rating will raise other subscribers' usage (i.e., they have "active expectations" about how individual usage reacts to zero-rating). The ISP then adopts zero-rating in order to increase expected surplus from subscribing; the concomitant increase in subscriber numbers then increases his profits.[4] Our modeling exercise also provides two further insights: First, it is consumers' active beliefs about *usage*, not the number of *subscribers*, that matters. Second, the ISP introduces zero-rating if and only if it would be socially optimal to do so (for given content).[5]

[2] This is similar in spirit to the "Family and Friends" offers of free voice calls to a limited set of numbers.

[3] 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.

[4] The Facebook-led initiative internet.org also uses zero-rating of certain services to increase participation, which has surpassed 100 million users in 2019 (https://info.internet.org/en/impact). As in other cases where zero-rating is not applied to competing services, this programme is controversial and has even been prohibited in India in 2016.

[5] In order to focus on the demand side, the content side is exogenous in our model. A full consideration of the effects of zero-rating on the economy must take into account its effects on content provision.

2 A Model of Zero-Rating with Network Effects

2.1 Setup

A monopolist ISP provides access to two types of content x and y, the former of which creates network externalities. Bandwidth usage of service x is $r > 0$ per unit, while that of service y is normalized to 1. The ISP offers a tariff (T, q, λ), where $T \geq 0$ is a fixed monthly fee, $q \geq 0$ is a data cap, and $\lambda \in [0, 1]$ is the degree of zero-rating. Individual data usage or capacity provided is $k = rx + y$, but only $\lambda rx + y$ are counted towards the data cap. Thus we have full zero-rating of service x for $\lambda = 0$, a joint data cap for $\lambda = 1$, and partial zero-rating otherwise. Capacity costs are $c > 0$ per unit.

Consumers' utility is given by $U(x, y, X) + \varepsilon$, where X is expected aggregate usage of service x, and ε is an individual taste parameter distributed according to a distribution H with density h. Thus consumers are homogeneous in usage but differ in their willingness to participate in the first place.[6] Utility U is twice continuously differentiable with $U_x > 0$ up to bliss points $\bar{x}(y, X) < \infty$ such that $U_x(\bar{x}(y, X), y, X) = 0$; $U_y > 0$; $U_X > 0$ (a participation network effect); U is strictly concave in (x, y), i.e. $U_{xx}, U_{yy} < 0$ and $U_{xx}U_{yy} - U_{xy}^2 > 0$; $U_{xy} \geq 0$, i.e. services are complements or independent;[7] $U_{xX} \gtreqqless 0$ (usage network effect if positive, or crowding out if negative), and $U_{yX} = 0$. Consumers' outside option has value 0, thus the number of consumers is given by

$$\alpha = \Pr(U + \varepsilon - T \geq 0) = 1 - H(-(U - T)) \equiv G(U - T),$$

with $G' = h$, and the ISP's profits are $\pi = \alpha(T - ck)$. Note that higher total consumption of service X, *ceteris paribus*, raises participation since $\alpha_X = hU_X > 0$.

The central assumption in this paper is that consumers have active (rational) expectations about individual usage $\bar{x}$, i.e. in particular they understand that whether the ISP offers zero-rating or not has a significant effect on *other consumers' usage,* and will take this into account in their subscription choice.[8] This implies that when choosing its tariff the ISP takes into account how this choice will influence customer expectations. On the other hand, it makes no difference for our argument whether consumers have active or passive expectations about the total number of customers, because price adjustments by the firm imply that the main channel driving total consumption is usage. For simplicity and in order to highlight the different role of usage expectations we opt for passive beliefs.

Thus the timing of the model is as follows: 1. Given beliefs $\bar{\alpha}$, the ISP chooses its tariff (T, q, λ); 2. Consumers observe the tariff and form beliefs $X = \bar{\alpha}\bar{x}$ about aggregate usage; 3. Consumers make subscription and usage decisions, and profits are realized.

[6] Assuming homogeneity in usage helps us concentrate on the issue of network effects, by eliminating the scope for price discrimination. With enough heterogeneity, zero-rating would also be used for the latter. Our results then imply that more zero-rating would be observed than that used for price-discrimination purposes.

[7] Though $rU_{xy} > U_{xx}$ would be enough for what follows.

[8] Under passive expectations, the ISP believes that its tariff choice does not influence consumer expectations about usage. Technically speaking, this amounts to applying the rational expectations condition of setting the latter equal to actual values only after the tariff choice.

2.2 Results

Given the tariff (T, q, λ) and expectations X, consumers choose their optimal consumption by solving $\max_x U(x, q - \lambda r x, X)$, with first-order condition $U_x = \lambda r U_y$ (the cap will be binding). Total capacity used is $k = q + (1-\lambda) r x$. This defines a bijection from (T, q, λ) to (T, k, λ), so we can solve the ISP's problem in terms of the latter.[9] Usage $\hat{x}(\lambda, k, \bar{\alpha})$ follows from[10]

$$U_x(\hat{x}, k - r\hat{x}, \bar{\alpha}\bar{x}) = \lambda r U_y(\hat{x}, k - r\hat{x}),$$

and active expectations $\bar{x} = \hat{x}$, resulting in surplus $\hat{U}(\lambda, k, \bar{\alpha}) \equiv U(\hat{x}, k - r\hat{x}, \bar{\alpha}\hat{x})$. Given $\bar{\alpha} > 0$, the ISP maximizes its profits as

$$\max_{\lambda, k, T} \pi = G\left(\hat{U}(\lambda, k, \bar{\alpha}) - T\right)(T - ck),$$

with derivatives[11]

$$\frac{\partial \pi}{\partial \lambda} = h\hat{U}_\lambda (T - ck), \quad \frac{\partial \pi}{\partial k} = h\hat{U}_k (T - ck) - \alpha c, \quad \frac{\partial \pi}{\partial T} = -h(T - ck) + \alpha.$$

From this is follows first that the profit-maximizing fixed fee is given by $T^* = ck + \frac{\alpha}{h}$ with profits $\pi^* = \alpha^2/h$, i.e. profits increase with the number of subscribers;[12] substituting T^* into the capacity condition shows that the optimal capacity equates benefits to cost, $\hat{U}_k = c$. As for the zero-rating decision, the marginal utility from raising λ is

$$\hat{U}_\lambda = (U_x - rU_y)\hat{x}_\lambda + \bar{\alpha}U_X \bar{x}_\lambda = -(\bar{\alpha}U_X - (1-\lambda) rU_y) rU_y/\Delta$$

with $\Delta \equiv (1-\lambda)(rU_{xy} - U_{xx}) - \lambda(U_{xx} - 2rU_{xy} + r^2 U_{yy}) - \bar{\alpha}U_{xX}$. We assume $\Delta > 0$, which, since the first two terms are positive, means that usage network effects cannot be explosively large.

Now our main result follows: A joint cap ($\lambda = 1$) is optimal if and only if the participation network effect U_X is zero, while full zero-rating ($\lambda = 0$) is chosen if

$$\bar{\alpha}U_X \geq rU_y, \qquad (1)$$

i.e. the benefits $\bar{\alpha}U_X$ due to higher X outweigh the opportunity cost rU_y of lower (distorted) consumption y.[13]

This makes clear that the decisive element for the ISP is the participation network effect U_X: Higher consumer expectations about other subscribers' usage drive up the willingness to subscribe in the first place. The ISP uses zero-rating to raise these expectations

[9] This leads to an equivalent system of first-order conditions but solves the problem faster.

[10] We suppress that last argument of U_y because $U_{yX} = 0$.

[11] We assume that the sufficient second-order conditions for a maximum hold.

[12] With active expectations about customer numbers we obtain the lower fixed fee $T = ck + \frac{\alpha}{h} - \alpha\hat{U}_\alpha$, with profits $\pi^* = \alpha^2(1/h - \hat{U}_\alpha)$, incorporating the trade-off of a lower fixed fee for more customers.

[13] An earlier version of this paper had $U(x, y, X) = u(x, y) + \beta X$, with participation network effect $U_X = \beta$ and usage network effect $U_{xX} = 0$. In the present setting it becomes clear that it is participation that counts. Note also that there is no contradiction between sufficiently strong participation and sufficiently weak usage network effects.

and therefore increase its subscriber numbers, and the data cap q is set equal to usage of the other service, $y = k - rx$.

On the other hand, zero-rating is only profit-maximizing if there is such a network effect or some other type of externality that raises profits. One such externality, as discussed in the previous literature, is payments from content providers to the ISP; if these are forbidden then in the absence of network effects zero-rating induces a distortion in individual usage towards the zero-rated service that actually reduces profits. This distortion also exists under network effects, but is outweighed by the increased subscriber base.

A decision by the ISP to zero-rate maximizes welfare. The latter is given by the aggregate surplus created by all subscriptions, i.e. $W = \int_{\varepsilon \geq -(\hat{U}-T)} (\hat{U} - ck + \varepsilon) dH(\varepsilon)$, with

$$\frac{\partial W}{\partial \lambda} = \left[-(-1)(\hat{U} - ck - (\hat{U} - T)) h(-(\hat{U} - T)) + \int_{\varepsilon \geq -(\hat{U}-T)} dH(\varepsilon) \right] \hat{U}_\lambda = 2\alpha \hat{U}_\lambda$$

at the equilibrium tariff. Therefore welfare is maximized at full zero-rating ($\lambda = 0$) if and only if it is profit-maximizing for the ISP ($\hat{U}_\lambda = 0$): there is no conflict between private and social incentives. The reason for this is that the ISP already maximizes social surplus by coordinating subscribers.

3 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 in order to raise the number of subscriptions. If chosen, zero-rating is socially optimal, thus no immediate market failure results. This indicates that a case-by-case review may be more adequate than a blanket prohibition of zero-rating offers.

References

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

[2] Gautier, A., and R. Somogyi (2018). Prioritization vs Zero-Rating: Discrimination on the Internet. CESifo Working Paper No. 7185.

[3] 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.

[4] Jaunaux, L., & M. Lebourges (2019). Zero rating and end-users' freedom of choice: an economic analysis, Digital Policy, Regulation and Governance, 21, 115-128.

[5] Jullien, B. and W. Sand-Zantman (2018). Internet Regulation, Two-Sided Pricing, and Sponsored Data, International Journal of Industrial Organization, 58, 31-62.

[6] 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.

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

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

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