Schnurr, Daniel; Wiewiorra, Lukas

Conference Paper
Bit-by-Bit Towards Unlimited: An Analysis of Zero Rating and Sponsored Data Practices of Internet Service Providers

29th European Regional Conference of the International Telecommunications Society (ITS): "Towards a digital future: Turning technology into markets?", Trento, Italy, 1st - 4th August 2018

Provided in Cooperation with:
International Telecommunications Society (ITS)


This Version is available at:
http://hdl.handle.net/10419/184965

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Bit-by-Bit Towards Unlimited: An Analysis of Zero Rating and Sponsored Data Practices of Internet Service Providers

Daniel Schnurr* Lukas Wiewiorra†

Abstract
Zero rating and sponsored data have recently been introduced as new tariff options in mobile telecommunications markets. Both practices exempt data traffic of specific Internet services from an Internet user’s download limit and thus allow for unlimited consumption of the respective content. Whereas network operators have advertised these options as consumer-friendly innovations, both practices have been criticized by net neutrality proponents and are currently under regulatory scrutiny in several countries. Based on a model of a monopolistic Internet access provider and two advertising-financed content providers, we investigate market outcomes under both zero rating and sponsored data and assess optimal pricing schemes, firms’ profits, consumer surplus and total welfare. We then analyze welfare effects if network operators are prohibited from offering service-specific exemptions from users’ data allowance and derive policy implications for the current regulatory debate on net neutrality.

Keywords: Internet access, data caps, zero rating, sponsored data, price discrimination, net neutrality, regulation, public policy
JEL classification: L51, L86, L96, D61

---

*Research Group Data Policies, University of Passau, Dr.-Hans-Kapfinger-Str. 12, 94032 Passau, Germany, phone: +49 851 509 2584, e-mail: daniel.schnurr@uni-passau.de.
†Department of Business Informatics and Information Economics, Goethe University Frankfurt, Germany, phone: +49 69 798 34708, e-mail: wiewiorra@econ.uni-frankfurt.de.
1 Introduction

In many countries Internet access tariffs include data allowances that define an upper limit to a user’s monthly Internet consumption in terms of aggregated traffic volume. Such data caps have been widely adopted in telecommunications markets, as they allow Internet service providers (ISPs) to price discriminate among users with varying valuations for Internet content and services. Data caps are most popular in mobile telecommunications markets, but have also been employed in fixed-line access networks (Brodkin, 2017).

In recent years, mobile network operators in the US and Europe have additionally introduced tariff options that exempt data traffic of specific content- and services providers (CSPs) from an Internet user’s data cap (see DotEcon, 2017, for an overview of zero rating practices). Such content-specific exemptions may either be offered by the ISP without any costs for the respective CSP, as in the case of zero rating, or only in return for a monetary payment or revenue share from the CSP, as in the case of sponsored data. In both cases, consumers can then receive unlimited data traffic from the respective CSP.

Whereas network operators have advertised zero rating and sponsored data as consumer-friendly innovations, both practices have been criticized by net neutrality proponents (Finley, 2015) and are currently under regulatory scrutiny in several countries (Krieger, 2017). Net neutrality refers to a regime where data traffic is supposed to be treated equally, independent from source, ownership or destination (Krämer et al., 2013; Greenstein et al., 2016). In particular, (exclusive) sponsored data agreements have been found to violate net neutrality rules, as they may discriminate between Internet content and give some services an undue competitive advantage (Federal Communications Commission, 2017). In contrast, zero rating arrangements that exempt entire content categories from a data allowance in a non-discriminatory manner are regularly viewed as net neutral (see, e.g.,

---

1 The terms zero rating and sponsored data are often used interchangeably, but in this article, we will distinguish both concepts according to the definitions outlined in this introduction.
the net neutrality guidelines adopted by BEREC, 2017, in the European Union). However, whether these diverging policies and regulatory decisions are actually supported from an economic perspective and whether they are likely to ultimately benefit consumers has so far not been addressed. Whereas the economic literature on net neutrality has extensively covered the issues of (paid) prioritization and termination fees, there is relatively few research on sponsored data and, especially, zero rating (see our survey of the related literature in Section 2).

In this study, we investigate zero rating and sponsored data practices in a unified framework. Based on a model of a monopolistic ISP and two advertising-financed CSPs, we investigate market outcomes and welfare measures in the case of both pricing schemes. More specifically, we explicitly consider tariffs with binding data caps on the user side as well as content-specific data costs. Next to an assessment of their relative profitability and market performance, this allows us to compare both practices to conventional (one-sided) pricing schemes in telecommunications markets, in particular, plain data caps and flat rates. By doing so, we can derive relevant policy implications with regard to current net neutrality obligations that prohibit specific practices, but allow for others.

We show that zero rating may hurt consumers if offered without compensation and voluntarily by the ISP, because it allows to extract more consumer surplus by the means of improved price discrimination. In contrast, sponsored data may increase consumer surplus, if it induces the ISP to offer data traffic exemptions in addition to plain data caps. However, content that is not zero-rated or sponsored may be worse off in this case compared to a situation with plain data caps. Moreover, zero rating and sponsored data may distort consumers’ consumption pattern to the detriment of such content providers even if they represent independent services, i.e., if they are not in direct competition with zero rated or sponsored services. These insights question the current consensus among policy makers and regulators in the European Union with respect to the practices that are allowed under
the codified net neutrality rules.

Whereas this paper focuses on zero rating and sponsoring of a single service or a services class, we intend to extend the scope of our analysis to competition between CSPs. In contrast to paid prioritization, which exerts a negative externality on competitors, zero rating and sponsored data exerts a positive externality, as it frees up additional space under the data cap for competing services. Moreover, our model setup allows us to consider heterogeneous data costs per content or service and thus to analyze zero rating and sponsored data decisions in the context of services with a varying degree of data-intensity (e.g., video and audio streaming services).

The remainder of this paper is structured as follows: In Section 2, we survey the related literature on zero rating and sponsored data. Section 3 introduces our model setup, followed by Section 4, that presents the analysis of four pricing schemes, including zero rating and sponsored data. Market outcomes and welfare results are reported in Section 5. Section 6 concludes and discusses limitations together with future work.

2 Related Literature

The recent emergence of zero rating and sponsored data practices in many telecommunications markets, together with the ongoing policy debate, has stimulated several research endeavours in this domain. In particular, these studies examine how the economic effects of those practices relate to the broader literature on net neutrality (Krämer et al., 2013; Greenstein et al., 2016). Whereas (paid) prioritization schemes and termination fees have been thoroughly studied in this context, Greenstein et al. (2016) highlight in their survey on net neutrality that whether zero rating practices “represent efficiency gains, or whether they unfairly tip the competitive landscape, raise the cost of rival services and provide a cause for regulator intervention” (p.132) remains among the unresolved policy issues.
Gautier and Somogyi (2018) compare zero rating and sponsored data to discriminatory prioritization of specific data traffic with respect to an ISP’s profits and the ensuing welfare outcomes. Whereas prioritization increases vertical quality differentiation between content providers, the former practices induce different financial costs for Internet users depending on the content that they consume. In this analysis different financial costs arise, because the ISP offers a two-part tariff that includes a marginal price instead of a data allowance. As both practices may create distortions in the content market, a monopolistic ISP can leverage them to extract additional profits from advertising-financed content providers.

Jullien and Sand-Zantman (2018) show that a monopolistic ISP can use sponsored data as a screening device to discriminate between high-benefit and low-benefit content providers that are advertising-financed. Whereas high-benefit content providers have an incentive to sponsor consumption in order to generate higher advertising revenues, low-benefit content providers prefer not to do so. In consequence, sponsored data induces efficient consumption of high-benefit content and ultimately increases total welfare if the share of high-benefit content providers is large. On the contrary, if the share of low-benefit content providers is large, one-sided pricing is more socially efficient.

In the study by Cho et al. (2016) sponsored data weakly dominates one-sided pricing of Internet users from the perspective of the ISP. However, sponsored data is only strictly preferred by the ISP if competition between content providers is sufficiently high. As consumers single-home between content providers, competition for advertisement impressions induces one or both content providers to sponsor data (which, in the latter case, may ultimately lead to a prisoner’s dilemma outcome). In line with Jullien and Sand-Zantman (2018), the authors show that the content provider, which is able to generate higher advertising revenues, will be more likely to subsidize content. From a welfare perspective, however, equal treatment of content providers (no sponsored data or non-discriminatory offers to all content providers) is always superior to an exclusive sponsoring agreement.
All of the above studies do not explicitly consider binding quantity limits for Internet users in the form of data caps. In Gautier and Somogyi (2018) and Cho et al. (2016) the ISP either offers a usage-based price or a two-part tariff with a fixed fee and a marginal fee for each unit of consumption. In Jullien and Sand-Zantman (2018) the data allowance is set to the unconstrained consumption level of each consumer. Yet, in practice, zero rating and sponsored data tariffs are predominantly marketed as exemptions from constraining download limits. As shown by Economides and Hermalin (2015) such data caps may represent an important strategic tool for Internet service providers with regard to content and services providers. In this study, we therefore take into account binding quantity limits, when we assess the effects of zero rating and sponsored data on upstream competition.

Moreover, with respect to the consumer side, data caps are a well-known instrument to implement second-degree price discrimination and are extensively used in this regard in mobile telecommunications markets. As shown by Inceoglu and Liu (2017), zero rating may be implemented by an ISP together with data caps in order to extract a higher surplus from consumers if they have heterogeneous preferences for content. In addition, the authors show that zero rating then generally leads to higher investments in network capacity. In this setting, the effect on welfare is ambiguous, depending on the competitive relationship among content providers and the degree of heterogeneity in consumers’ valuation.

Finally, Somogyi (2017) abstracts from price discrimination on the consumer side, but considers binding data caps in the context of a capacity-constrained ISP that offers sponsored data to content providers. In this setting, the ISP is willing to offer sponsored data to content providers that are either very unattractive or very attractive to consumers, but not to content providers that offer an intermediate quality. Whereas sponsoring of high-quality content increases total welfare, sponsoring of low-quality content leads to a welfare decrease. In general, consumers may be negatively affected by sponsored data if the negative externality from congestion outweighs utility gains from increased consumption.
In this study, we consider both zero rating and sponsored data in a single analytic framework. Based on this framework, we identify the conditions for which a monopoly ISP is willing to introduce zero rating for a single service or a specific class of content in addition to plain data caps. Moreover, we characterize the conditions when a CSP would be willing to engage in a sponsored data deal, if offered by the ISP. Contrasting zero rating and sponsored data as well as comparing those practices to established tariff options in mobile telecommunications markets allows us to analyze welfare effects for the different industry stakeholders. Specifically, we can thus inform the regulatory debate on whether these practices should be allowed or prohibited on the basis of net neutrality concerns. As we explicitly analyze data cap sizes as well as consumer surplus, we can test whether the former is a good proxy for the latter, as cap sizes are frequently discussed as a reference point for consumer surplus in practice. Moreover, we can evaluate the general regulatory consensus that sponsored data should be met with stronger scrutiny than zero rating, as the latter is frequently viewed as the more consumer-friendly practice.

3 The model

Internet service provider (ISP): We consider the case of a monopolistic ISP that intermediates data traffic between Internet users and CSPs. Thus, the ISP represents a two-sided platform that, in principle, can charge both market sides. However, as discussed in the preceding sections, net neutrality regulation may constrain the ISP in its ability to charge CSPs. We will therefore not consider pricing mechanisms that establish direct termination fees for CSPs (see, e.g., Economides and Tåg, 2012, on this issue). Instead, we focus on pricing strategies that exempt data traffic of specific services and content from an Internet user’s data allowance. Whereas such zero rating may be offered by the ISP without an explicit fee for any market side, those exemptions may also come in the form
of sponsored data. In the latter case, we will assume that the ISP receives a share $s$ of the CSP’s revenue in return for exempting the respective traffic from a user’s data cap. Moreover, we assume that the ISP does not face any capacity constraints and normalize marginal costs of providing Internet access to zero.

**Content- and services providers (CSPs):** In line with the related literature, we focus on advertising-financed CSPs that offer their content and services for free to consumers. We normalize the costs of providing content of each CP to zero and assume that ads can be displayed at zero marginal costs. Thus, the profit function of CSP $i \in [A, B]$ is given by $
pi_i := (1 - s)r_q$, where $r$ represents the advertising revenue of the CSP if an Internet user consumes one unit of its content and $s$ represents the revenue share that a CSP may be required to pay in the case of sponsored data. In the case of zero rating, there is no revenue sharing, i.e., $s = 0$, and the CSP’s profit function reduces to $
pi_i := r_q$. For simplicity, we will assume that CSPs are symmetric, i.e., they generate the same advertising revenues for one unit of consumption $q_i$.

**Internet users:** There is a unit mass of Internet users that consists of two heterogeneous consumer types. For each consumer type $k$, we consider a representative consumer à la Singh and Vives (1984) that receives utility $U_k := \alpha_A q_A + \alpha_B q_B - \frac{1}{2}(\beta_A q_A^2 + \beta_B q_B^2 + 2\gamma q_A q_B)$ when consuming $q_i$ units ($i = A, B$) of the respective content and services. Parameters $\alpha_i$ and $\beta_i$ specify the precise form of the quadratic utility function and allow us to consider various relative valuations between the offered content and services. The relationship between CSP $A$ and $B$ as viewed by consumers is measured by the substitution parameter $\gamma$. If $\gamma < 0$ services and content are complementary, if $\gamma = 0$ they are independent of one another, and if $\gamma > 0$ they are substitutes.

Consumption of CSP $i$’s content and services generates data traffic $D_i = d_i q_i$, where $d_i$ denotes the amount of data traffic volume that is associated with the consumption of one
unit of content. Cumulative data traffic counts against a user’s data cap $C_i$, if the user has signed up for a tariff that includes a data allowance and if the respective content or service is not zero rated or sponsored. Thus, CSPs may not only be differentiated with respect to the utility that they offer to consumers, but may also be heterogeneous with respect to the data costs that they impose on consumers. This allows us to capture content classes that differ in their data intensity, such as video streaming and messaging services as well as relative advantages in data compression techniques of services in the same content class.

**Consumer’s preferences for content and services:** Throughout this study, we will assume that the population in our model consists of two consumer segments $k \in [\text{H}, \text{L}]$, where $\text{H}$ denotes a segment of high-value consumers and $\text{L}$ denotes a segment of low-value consumers. As a reference point we first fix the preference for CSP $\text{A}$ in the population and vary the preference for CSP $\text{B}$ in order to construct a spectrum of relevant preference scenarios. To that end, we fix the preference of consumer group $\text{H}$ for CSP $\text{A}$ on a high level ($\alpha_{\text{H}, \text{A}}$) and the preference of consumer group $\text{L}$ for both CSPs on a low level ($\alpha_{\text{L}}$).

Given this preference order for CSP $\text{A}$, we vary the preference for CSP $\text{B}$ of the high-value consumers, as depicted in Figure 1: This distribution of tastes implies that a consumer that likes CSP $\text{A}$ will also like CSP $\text{B}$, and vice versa. Note that this distribution of preferences results in a maximum asymmetry in data traffic between consumer segments if $\alpha_{\text{H}, \text{B}} = \alpha_{\text{H}, \text{A}}$.

![Figure 1: Heterogeneity of consumers.](image-url)
Varying the degree of heterogeneity in this way allow us to analyze the profitability of zero rating and sponsored data if one consumer group is either more or less in favor of both CSPs. This differentiation is of fundamental importance for the following analysis, because higher values of $\alpha_{HB}$ make both consumer groups more distinct from each other with respect to data consumption and valuation for content.\(^2\)

**Pricing schemes and timing:** We evaluate zero rating and sponsored data on the basis of firms’ profits and prices as well as consumer surplus and total welfare. We then compare both practices to two-part tariffs with a plain data cap and a fixed subscription fee, which is currently the predominant tariff option in mobile telecommunications markets. As a further benchmark, we consider a flat rate offer which represents the most popular pricing scheme in fixed telecommunications markets, but has also been introduced in US and European mobile markets.\(^3\) We will at first analyze optimal price setting of the ISP for each of the four pricing schemes separately in Section 4 and then compare equilibrium profits and market outcomes in Section 5. Thereby, we consider the following timing of firms’ and users’ actions:

*Stage I:* The ISP decides whether to make a sponsored data offer to one of the CSPs.

*Stage II:* The respective CSP decides whether to sponsor data, if offered by the ISP.

*Stage 1:* The ISP decides which tariff(s) and fee(s) to offer to users.

*Stage 2:* Internet users decide whether to sign up with the ISP and which tariff to choose.

Subscribed users consume services of CSP A and/or CSP B.

Note that Stage I and Stage II are only relevant in the case of sponsored data, but Stage 1 and Stage 2 are relevant for all types of tariffs.

\(^2\)Note that we introduce *intra-group heterogeneity* for high-value consumers, whereas we homogenize low-value consumers. We could have also assumed a reverse setting, i.e., homogeneity of high-value consumers and heterogeneity with respect to low-value consumers.

\(^3\)See, for example, the recent announcement of the German incumbent Deutsche Telekom (2018) to introduce an unlimited mobile tariff plan.
4 Analysis of pricing schemes

Next, we solve for the optimal tariff menu offered by the ISP for each of the four pricing schemes by backward induction. Optimal tariff properties, such as prices, cap sizes as well the respective firms’ profit and users’ consumption in equilibrium are then compared and discussed in Section 5.

4.1 Flat rate

In the case of a flat rate, the ISP offers a single tariff with a subscription fee $F = F^{FR}$ that allows all subscribers to consume unlimited traffic of both CSPs. The ISP then makes the following profit:

$$\pi^{FR}_I := \begin{cases} 
\theta F & \text{if only high-type consumers subscribe with the ISP,} \\
F & \text{if all consumers subscribe with the ISP.}
\end{cases}$$

**Consumers:** In Stage 3, consumers of type $k \in [H, L]$ who sign up for the flat rate choose unconstrained consumption levels $q_k^*$, maximizing their net utility:

$$\max_{q_{kA}, q_{kB}} U_k(q_{kA}, q_{kB}) - F.$$ 

Facing two heterogeneous consumer groups the ISP, in Stage 1, compares the profit with a flat rate for all consumers at a low subscription fee to the profit with a flat rate for high-type consumers only at a high subscription fee. In the former case, the subscription fee is set such that the entire surplus of low-type consumers is extracted by the ISP, whereas high-type consumers receive a positive net utility. In the latter case, the surplus of high-type consumers is fully extracted through the subscription fee, whereas low-type consumers choose not to sign up with the ISP. The consumption level of high-type consumers is then
given by:

\[
q_{HA}^{FR} = \frac{\beta \alpha_{HA} - \gamma \alpha_{HB}}{\beta^2 - \gamma^2},
\]

\[
q_{HB}^{FR} = \frac{\beta \alpha_{HB} - \gamma \alpha_{HA}}{\beta^2 - \gamma^2}.
\]

The consumption level of subscribers with a low valuation is given by:

\[
q_{L_i}^{FR} = \frac{(\beta - \gamma)\alpha_L}{\beta^2 - \gamma^2}.
\]

**Internet service provider:** Recall that $\theta$ defines the share of high-value consumers in the market. Moreover, we denote the threshold level of high-value consumers in the market up to which the ISP will serve both types of consumers by $\tilde{\theta}^{FR}$. If the share of high-value consumers exceeds this threshold, the ISP will exclude low-value consumers and cater only to high types:

\[
\tilde{\theta}^{FR} = \frac{2\alpha_L^2(\beta - \gamma)}{(\beta \alpha_{HA}^2 + \beta \alpha_{HB}^2 - 2\gamma \alpha_{HA} \alpha_{HB})}.
\]

In consequence, the optimal access fee of the ISP depends on the share of high-type consumers in the market, because exclusion becomes less costly (in terms of opportunity costs) if the share is rising. The ISP can extract the complete consumer surplus if only high-value consumers are served. If the ISP caters to both consumer groups the ISP can only fully extract the surplus from low-value consumers, whereas high-value consumers retain a surplus. The optimal schedule of access fee’s of the ISP is therefore given by:

\[
F^{FR} = \begin{cases} 
\frac{\alpha_L^2}{\beta + \gamma} & \text{if } 0 \leq \theta < \tilde{\theta}^{FR}, \\
1 \left(\frac{\alpha_{HA}^2 + \alpha_{HB}^2}{\beta^2 - \gamma^2}\right) - 2\gamma \alpha_{HA} \alpha_{HB} & \text{if } \tilde{\theta}^{FR} \leq \theta < 1.
\end{cases}
\]
4.2 Data Cap

In the case of plain data caps, the ISP may offer two tariffs to price discriminate between consumers with a high and low valuation for content. Specifically, the ISP offers a portfolio $T$ of two-part tariffs with a data cap $C$ and a fixed subscription fee $F$, i.e., $T^{DC} = \{(F_H, C_H); (F_L, C_L)\}$.

**Consumers:** To obtain the optimal portfolio of tariffs we first derive the consumer’s best response for any given data cap. Note that, compared to the case of a flat rate, a consumer’s optimization problem is now subject to additional constraints due to the quantity restriction. In consequence, consumers are rationed in their maximum consumption of Internet traffic, which is similar to a budget constraint with exogenous prices (i.e., data costs). Formally, a consumer solves the following optimization problem

$$\max_{q_A, q_B} \alpha_A q_A(C) + \alpha_B q_B(C) - \frac{1}{2} \left( \beta_A q_A(C)^2 + \beta_B q_B(C)^2 + 2\gamma q_A(C)q_B(C) \right) - F$$

subject to $(d_A q_A(C) + d_B q_B(C) \leq C)$.

From a consumer’s perspective the tariff portfolio $T^{DC}$ is given. Therefore, consumers compare their expected utility from consuming Internet services under both tariffs in the portfolio and subsequently i) opt for the tariff with the higher expected net-utility or ii) choose the outside option and do not sign up with the ISP. The optimal demands of consumers with a high valuation, who sign up with the ISP, are given by

$$q^{DC}_{H_A} = \frac{2\alpha_{H_A} + C(\beta d_A - \gamma d_B) - \alpha_{H_B} d_A d_B}{\beta(d_A^2 + d_B^2) - 2d_A d_B \gamma},$$

$$q^{DC}_{H_B} = \frac{2\alpha_{H_B} + C(\beta d_B - \gamma d_A) - \alpha_{H_A} d_A d_B}{\beta(d_A^2 + d_B^2) - 2d_A d_B \gamma}.$$
The demands of consumers with a low valuation, who sign up with the ISP, are given by
\[ q^{DC}_{L_i} = \frac{2\alpha_L + C(\beta d_A - \gamma d_B) - \alpha_L d_A d_B}{\beta(d_A^2 + d_B^2) - 2d_A d_B \gamma}. \]

**Internet Service Provider** In order to successfully discriminate between both segments of consumers via self-selection the ISP has to solve the following optimization problem with respect to incentive and participation constraints for both consumer segments:

\[
\begin{align*}
\max_{F_H, F_L, C_H, C_L} \quad & \Pi_I = \theta F_H + (1 - \theta) F_L \\
\text{s.t.} \quad & U_H [q^{DC}_{H_A}(C_H), q^{DC}_{H_B}(C_H)] - F_H \geq U_H [q^{DC}_{H_A}(C_L), q^{DC}_{H_B}(C_L)] - F_L \quad (IC_H) \\
& U_L [q^{DC}_{L_A}(C_L), q^{DC}_{L_B}(C_L)] - F_L \geq U_L [q^{DC}_{L_A}(C_H), q^{DC}_{L_B}(C_H)] - F_H \quad (IC_L) \\
& U_H [q^{DC}_{H_A}(C_H), q^{DC}_{H_B}(C_H)] - F_H \geq 0 \quad (PC_H) \\
& U_L [q^{DC}_{L_A}(C_L), q^{DC}_{L_B}(C_L)] - F_L \geq 0 \quad (PC_L)
\end{align*}
\]

In line with the general properties of optimal second-degree price discrimination, the following pairs of prices and quantity limits ensure that: i) Consumption of high-value consumers is not distorted, i.e., the data cap designed for high types is non-binding and, thus, they enjoy unconstrained consumption. However, the ISP might have to leave high-value consumers with a surplus in order to satisfy the incentive constraint. ii) Surplus from low-value consumers is always fully extracted by the ISP and their consumption can be restricted to render the alternative for high-value consumers (who may pretend to be low-type consumers) less attractive in order to satisfy the incentive constraint.

To calculate the optimal subscription fee for each consumer segment, we have to dif-
ferentiate between several cases depending on the optimal cap size for low-type consumers \( C_L^{DC} \). Specifically, the following cases have to be considered (assuming \( d_A \leq d_B \)):

1. The ISP offers a portfolio of tariffs \( T_1^{DC} \) that would allow high-value consumers to consume both services (\( A \& B \)) if they pretend to be low-type consumers. At the same time, low-value consumers choose to consume both services under the tariff designed for them. This case holds true as long as the share of high-type consumers is not too high, i.e., if \( 0 \leq \theta \leq \hat{\theta}_1^{DC} \), with

\[
\hat{\theta}_1^{DC} := \frac{(\gamma - \beta)((\alpha_L - \alpha_B) d^2_A - \beta \alpha_B) + ((\alpha_{A_H} - \alpha_L) \beta + \gamma(\alpha_L + \alpha_A H)) d_B d_A - \beta d_B^2 \alpha_L}{(2d_A^2 d_B \gamma - \beta d_A^2 - \beta d_B^2)(\gamma \alpha_{A_H} - \beta \alpha_B)}. 
\]

2. The ISP offers a portfolio of tariffs \( T_2^{DC} \) that would allow high types to consume only their preferred service (i.e., CSP \( A \)) if they pretend to be low-value consumers. Low types still consume both services under the tariff designed for them. This case holds true if \( \hat{\theta}_1^{DC} \leq \theta \leq \hat{\theta}_2^{DC} \), with

\[
\hat{\theta}_2^{DC} := \frac{(\gamma - \beta) d_A \alpha_L}{(((d_B - d_A) \alpha_L - d_B \alpha_A H) \beta + d_A \gamma \alpha_A H)}. 
\]

3. The ISP offers a portfolio of tariffs \( T_3^{DC} \) if, in addition to (lying) high-value consumers, low-value consumers also consume only one of both services under the tariff designed for them. This is never the case for symmetric data costs, i.e., \( d_A = d_B \), because then low-type consumers value both services equally (as \( \alpha_A = \alpha_B = \alpha_L \)). However, if data costs are asymmetric, i.e., \( d_A < d_B \), low-value consumers may only consume CSP \( A \). This case holds true if \( \hat{\theta}_2^{DC} \leq \theta \leq \hat{\theta}_3^{DC} \), with

\[
\hat{\theta}_3^{DC} = \frac{\alpha_L}{\alpha_{A_H}}. 
\]

4. Otherwise, for \( \hat{\theta}_3^{DC} \leq \theta \leq 1 \), it is profitable for the ISP to exclude low-value consumers from the market by setting a cap size of zero, i.e., \( C_L = 0 \). In consequence, the results
from the flat-rate scenario apply and only high-types are served.

Whereas the ISP always finds it optimal to grant high-value consumer unconstrained consumption, low-value consumers will generally be constrained. The optimal size of the data cap for low-type consumers is given by:

\[
C_{L}^{DC} = \begin{cases} 
\frac{d_{A}((\theta \alpha_{A,H} - \alpha_{L})\beta + \gamma \alpha_{L})d_{A} - (2\left(\frac{1}{2}\beta \alpha_{L} + \gamma (\theta \alpha_{A,H} - \frac{1}{2} \alpha_{L}))d_{B}d_{A} + \beta d_{B}^{2} \theta \alpha_{A,H})}{((-1 + \theta) \beta^2 + \gamma^2) d_{A} - 2\beta d_{A}d_{B} \gamma \theta + \beta^2 d_{B}^2 \theta} & \text{if } 0 \leq \theta < \hat{\theta}_1^{DC} \\
\frac{(d_{A} \alpha_{A,H} + d_{B} \alpha_{B,H}) \beta + \gamma (d_{A} \alpha_{B,H} - d_{B} \alpha_{A,H}) \theta - (d_{A} + d_{B}) (\beta - \gamma) \alpha_{L}}{(1 - \theta) (\beta^2 - \gamma^2)} & \text{if } \hat{\theta}_1^{DC} \leq \theta < \hat{\theta}_2^{DC} \\
\frac{\alpha_{L} - \theta \alpha_{A,H} d_{A}}{(1 - \theta) \beta} & \text{if } \hat{\theta}_2^{DC} \leq \theta < \hat{\theta}_3^{DC} \\
0 & \text{if } \hat{\theta}_3^{DC} \leq \theta \leq 1
\end{cases}
\]

Figure 2a illustrates optimal cap sizes, fixed fees, and the ISP’s profit in the case of data cap tariffs without a zero rating or sponsored data option. Depicting the ISP’s equilibrium choices across the entire population spectrum, the figure visualizes the impact of the relative share of consumer segments. For low values of \( \theta \), the ISP finds it optimal to offer low-value consumers a positive data allowance and extract the entire surplus that those consumers receive from content consumption under the data cap. As the share of high-value consumers increases, the cap size \( C_{L}^{DC} \) decreases, which also results in a lower subscription fee \( F_{L}^{DC} \). On the other hand, high-value consumers always receive a data allowance that satisfies their unconstrained consumption of both CSPs. Whereas for low values of \( \theta \), the ISP must leave high-value consumers with a higher surplus in order to satisfy the incentive constraint, the ISP increases the subscription fee \( F_{H}^{DC} \) for higher values of \( \theta \). This is, because the outside option for high-value consumers becomes less attractive as the data cap in the tariff designed for low-value consumers \( C_{L}^{DC} \) is reduced. Eventually, for \( \theta > \hat{\theta}_3^{DC} \), it is profit-maximizing for the ISP to reduce the data cap \( C_{L}^{DC} \) to zero and thus to exclude low-value consumers from the market. This allows the ISP to offer a flat rate to high-type consumers and extract their entire surplus through the fixed
fee, as their incentive constraint now equals the participation constraint.

Figure 2: Optimal tariffs with plain data caps and zero rating. (For: $\alpha_{AH} = 2, \alpha_{BH} = 1.5, \alpha_L = \beta = d_A = d_B = 1, \gamma = 0$).

4.3 Zero-Rating

As we focus on exclusive zero rating offers, the ISP has generally the option to either zero rate CSP A or CSP B. For ease of presentation, we will assume for now that the ISP zero rates data traffic of CSP B.

Consumers: Facing a tariff portfolio $T^{ZR} = \{F_H, C_H, ZR_B\}; \{F_L, C_L, ZR_B\}$, where each tariff specifies a data cap $C$, a fixed fee $F$ and zero rating of CSP B, consumers in Stage 2 solve the following constrained optimization problem:

$$\max_{q_A, q_B} U_k(q_A, q_B) - F \quad s.t. \quad d_A q_A \leq C,$$
which yields optimal consumption quantities

\[
q_{kA}^{ZR} = \frac{C}{d_A}, \\
q_{Hb}^{ZR} = \frac{d_A \alpha_{Hb} - \gamma C}{\beta d_A}, \\
q_{LB}^{ZR} = \frac{d_A \alpha_{LB} - \gamma C}{\beta d_A}.
\]

**Internet service provider:** As in the case of plain data caps, the ISP, in Stage 1, designs its optimal menu of tariffs for high-type and low-type consumer segments \(T^{ZR} = \{(F_H^{ZR}, C_H^{ZR}, ZR_B); (F_L^{ZR}, C_L^{ZR}, ZR_B)\}\), such that subscription fees maximize its overall profit and at the same time fixed fees and cap sizes satisfy consumers’ incentive and participation constraints. Formally, the ISP solves the following optimization program:

\[
\max_{F_H, F_L, C_H, C_L} \pi_I = \theta F_H + (1 - \theta) F_L \\
\text{s.t.} \\
U_H [q_{HA}(C_H), q_{HB}^{ZR}(C_H)] - F_H \geq U_H [q_{HA}^{ZR}(C_L), q_{HB}^{ZR}(C_L)] - F_L \quad (IC_H) \\
U_L [q_{LA}^{ZR}(C_L), q_{LB}^{ZR}(C_L)] - F_L \geq U_L [q_{LA}^{ZR}(C_H), q_{LB}^{ZR}(C_H)] - F_H \quad (IC_L) \\
U_H [q_{HA}^{ZR}(C_H), q_{HB}^{ZR}(C_H)] \geq F_H \quad (PC_H) \\
U_L [q_{LA}^{ZR}(C_L), q_{LB}^{ZR}(C_L)] \geq F_L \quad (PC_L)
\]

Optimally, the ISP will design the tariff for high-type consumers such that they can consume their unconstrained consumption level \((q_{HA}^{ZR}, q_{HB}^{ZR})\). The data cap size in the tariff for low-type consumers is chosen such that the incentive constraint for high-type consumers is satisfied. Thus, given that the ISP does not exclude any consumer segment
through high fixed fees, optimal cap sizes with zero rating are given by

\[ C_{ZR}^{H} = d_A q_{HA} = \frac{d_A (\beta \alpha_{HA} - \gamma \alpha_{HB})}{\beta^2 - \gamma^2}, \]

\[ C_{ZR}^{L} = \begin{cases} 
\frac{d_A (-\beta \alpha_{HA} + \gamma \alpha_{HB}) \theta - \alpha_L (\gamma - \beta)}{(\theta - 1)(\gamma^2 - \beta^2)} & \text{if } 0 \leq \theta < \tilde{\theta}_{ZR}, \\
0 & \text{if } \tilde{\theta}_{ZR} \leq \theta < 1,
\end{cases} \]

and \( \tilde{\theta}_{ZR} := \frac{\alpha_L (\beta - \gamma)}{\beta \alpha_{HA} - \gamma \alpha_{HB}}. \)

Optimal fixed fees for high-type and low-type tariffs are given by \( F_{ZR}^{L} \) and \( F_{ZR}^{H} \), respectively (see Appendix A). Figure 2b illustrates the optimal menu of tariffs \( T_{ZR} \) and the resulting ISP profit depending on the distribution of consumer types. As for the case of plain data cap tariffs, the ISP is willing to offer low-type consumers a tariff with a positive cap size if the share of high-type consumers \( \theta \) is sufficiently small, i.e., if \( \theta < \tilde{\theta}_{ZR} \). The cap size \( C_{ZR}^{L} \) is decreasing in the share \( \theta \) as it becomes more profitable to increase \( F_{ZR}^{H} \) and satisfy the incentive constraint for high-value consumers by the means of a worse outside option. In order to still satisfy the participation constraint for low-type consumers, the respective subscription fee decreases accordingly for \( \theta < \tilde{\theta}_{ZR} \). In contrast to plain data cap tariffs, low-type consumers have a positive willingness to pay even if the data cap size is reduced to zero (i.e., for \( \theta \geq \theta_{ZR} \)), because they are still able to consume zero rated content of CSP B. In turn, this implies that, as long as low-type consumers are not excluded from the market through a high subscription fee, the ISP cannot extract the entire utility of high-type consumers and must leave them with a positive surplus in order to satisfy the incentive constraint. In Appendix A we show that the ISP may in some cases indeed have an incentive to exclude low-type consumers and derive the respective thresholds. Naturally, optimal cap sizes in the zero rating tariff are generally lower than in plain data cap tariffs, as traffic of CSP B consumption does not count against the data allowance. This is
best illustrated by the data caps for high-type consumers $C_H^{DC}$ and $C_H^{ZR}$, which both allow high-type consumers to consumer their unconstrained consumption of content from both CSPs.

![Figure 3: ISP’s profit with zero rating depending on the zero rated CSP. (For: $\alpha_{HA} = 2, \alpha_{BH} = 1.5, \alpha_L = 1, \beta = d_A = d_B = 1, \gamma = 0$).](image)

**Zero rating CSP A:** So far, we have assumed that the ISP will zero rate CSP $B$, i.e., the (on average) less-valued of both CSPs by high-type consumers. If instead CSP $A$ is zero rated, the presented analysis can be conducted analogously. Appendix A.2 denotes the equilibrium profit of the ISP $\pi_I^{ZR_A}$ in that case, assuming no exclusion of low-type consumers. Figure 3 depicts equilibrium profits for both alternative zero rating agreements and shows that for the specific parameter values, zero rating of CSP $B$ generates a higher profit for the ISP than zero rating of CSP $A$. In fact, the ISP is only indifferent between zero rating any of the two services if they are valued equally by high-value consumers, i.e., $\alpha_{HA} = \alpha_{HB}$, or if exclusion of low-type consumers is profitable. Otherwise, the difference in profits resulting from these alternative zero rating agreements $\Delta \pi_I = \pi_I^{ZR_B} - \pi_I^{ZR_A}$ increases
as the intra-group heterogeneity of content valuations among high-type consumers grows, i.e., as $\alpha_{H_b}$ decreases within interval between the reference points $[\alpha_L, \alpha_{H_A}]$.

**Result 1.** Zero rating of CSP $B$ weakly dominates zero rating of CSP $A$ for the ISP. If $\alpha_{A_H} > \alpha_{B_H}$ and low-type consumers are not excluded from the market, zero rating of CSP $B$ is strictly more profitable.

### 4.4 Sponsored Data

Figure 4 illustrates the timing of actions if the ISP considers a sponsored data tariff. At each stage, the respective actor weighs its profit in the case of sponsored data against the possible outside options. Sponsored data may increase the joint surplus of the ISP and the sponsoring CSP. In particular, a ISP may be able to extract revenue from a CSP through a positive revenue share $s$ and, in return, the CSP receives unconstrained user traffic in cases where the ISP would originally prefer plain data caps to zero rating.

If the ISP and CSP $B$ agree on a sponsored data deal, consumers’ equilibrium demand for CSPs in Stage 2 will be the same as in the zero rating tariff, i.e., $(q_{ZA}^{ZR}, q_{ZB}^{ZR})$ for both consumer types $k$. As shown in Section 4.3, the ISP generally prefers to zero rate CSP $B$. Therefore, we will assume that the ISP also makes an exclusive offer to the same CSP in the case of sponsored data.

In Stage 1, the ISP may either (i) implement zero rating tariffs (i.e., $T^{ZR}$) in case that the CSP has previously agreed to sponsor data or (ii) choose the tariff type that maximizes profits on the consumer side in case that the CSP has rejected the sponsored data offer. In the latter case, the ISP may either offer an optimal tariff menu of plain data caps ($T^{DC}$) or an optimal zero rating tariff menu ($T^{ZR}$). As shown for both tariff types (see, e.g., Figure 2), such optimal tariff plans may also come in the form of a flat rate for high-type consumers, excluding low-value consumers.
If the CSP receives a sponsored data offer from the ISP in Stage II, it will agree to this offer and the demanded revenue share $s$, if sponsoring increases its profit relative to the anticipated outside option. Therefore, the CSP has no incentive to accept any positive revenue share ($s > 0$) if the ISP will implement zero rating in the following stage anyway. In contrast, the CSP may be willing to sponsor data if the ISP will otherwise implement plain data caps. Specifically, the CSP agrees to sponsor data in this case if $\pi_i^{SD}(s) \geq \pi_i^{DC}$.

In other words, the CSP is willing to accept at most a maximum revenue share of $\bar{s}$, which leaves it indifferent between the outcome with sponsored data and the outcome with

---

**Figure 4**: Timing of decisions if the ISP considers a sponsored data tariff.
plain data caps. Therefore, we solve for \( \bar{s} \), which satisfies

\[
\pi^{SD}_i(\bar{s}) = \pi^D_i
\]

\[
\Leftrightarrow (1 - \bar{s})\pi_i(q_{HR}^Z + q_{LR}^Z) = \pi_i(q_{HR}^D + q_{LR}^D).
\]

Otherwise, if zero rating represents a second best from the ISP perspective, a threat to deviate to plain data caps is not credible and the CSP will therefore reject any offer with \( s > 0 \) by the ISP. In other words, if the ISP is willing to sponsor (i.e., zero rate) the CSP itself, that CSP has obviously no incentive to pay for sponsoring. Note that under the assumed timing, the ISP is not able to credibly commit to offer sponsored data as the exclusive tariff type. Arguably, this reflects the situation in practice, where ISPs struggle to commit to a two-sided pricing scheme, that would require them to forego profits on the consumer side.\(^4\)

**Result 2.** Sponsored data agreements weakly increase the revenues of the sponsoring CSP. However, the ISP is able to extract all additional revenues (in excess of revenues from plain data caps) from the CSP.

In Stage I, the ISP decides about a sponsored data offer based on its knowledge of the maximum revenue share \( \bar{s} \) that the CSP would still be willing to accept. Obviously, any sponsored data agreement with a positive \( s \) weakly dominates a zero rating tariff with the same CSP from the perspective of the ISP. Therefore, the ISP will make a sponsored data offer if and only if the anticipated profits outweigh the profit under a plain data cap tariff. Formally, the ISP offers a sponsored data deal with revenue share \( s = \bar{s} \) to CSP B in Stage I if and only if

\[
\pi^{SD}_i(\bar{s}) := \pi^Z_i + \bar{s}\pi^Z_B > \pi^D_i.
\]

\(^4\)See for example the case of Proximus in Belgium as described in Telecompaper (2018).
Whether this relation can be satisfied depends on the level of intra-group heterogeneity within the high-value consumer segment, i.e., the difference of $\alpha_{HA}$ and $\alpha_{HB}$. As can be seen in Figure 5a, the ISP cannot profitably offer sponsoring to CSP $B$ if $\alpha_{HB}$ is too different from $\alpha_{HA}$ (here $\alpha_{HA} = 2$, $\alpha_{HB} = 1.1$). In that case, the ISP finds zero rating profitable and would choose to exempt the CSP from consumers’ data allowance anyway. Therefore, in this case, any sponsored data offer with a positive revenue share $s$ is not credible. As a consequence sponsored data yields the same results as zero rating, because the only feasible revenue share is $s = 0$. In contrast, the ISP is able to extract additional revenues from the CSP side of the market by the means of sponsored data, if intra-group heterogeneity is low, i.e., if $\alpha_{HB}$ is close to $\alpha_{HA}$. For example, this is the case for $\alpha_{HA} = 2$ and $\alpha_{HB} = 1.6$, as depicted in Figure 5b.

![Figure 5: Sponsored Data ISP](image)

(a) ISP profit for $\alpha_{BH} = 1.1$. (b) ISP profit for $\alpha_{BH} = 1.6$.

**Result 3.** Sponsored data agreements can only be profitably offered by a monopolistic ISP if otherwise plain data caps would lead to higher profits than zero rating. If instead the ISP
prefers zero rating to plain data caps, the CSP will reject any offer with $s > 0$ and receive benefits from sponsoring for free.

Note that the assumed timing implies a higher bargaining power for the ISP. If we would assume a reverse order of actions (i.e., the CSP now makes a sponsoring offer to the ISP in the first stage) the equilibrium revenue share would generally be lower but still not always equal zero. The intuition for that result is as follows: In the cases where the ISP prefers plain data caps to zero rating, sponsored data will generally reduce the ISP’s ability to price discriminate among consumer groups. To that end, the CSP has to make an offer that (at least) compensates the ISP for its forgone revenues from better price discrimination. This condition would imply a lower bound ($s = \bar{s}$) for offers made by the CSP to the ISP to be accepted. Ultimately, however, the different revenue shares would merely imply a welfare neutral shift of rents between both firms.

5 Market outcomes and welfare analysis

Having derived equilibrium values for the different pricing schemes, we can now compare market outcomes and performance metrics of tariff portfolios across different scenarios.

ISP profit: Figure 6 depicts the ISP’s optimal profit for the three one-sided pricing strategies (flat rate, data cap and zero rating) for two sets of parameter values. Whereas the left panel displays one extreme case, where valuations of high-type consumers for content are identical (i.e., $\alpha_{HA} = \alpha_{BH}$), the right panel presents the opposite extreme case, where high-type consumers only differ in their valuation for content $A$ from low-type consumers (i.e., $\alpha_{HB} = \alpha_{LB}$). Naturally, profits are generally lower in the latter case if the share of high-value consumers $\theta$ is high, because there is less (high-type) consumer surplus in the market, as denoted by the variable $U_{max}$. Moreover, data caps are clearly a profitable
pricing strategy for the ISP compared to a flat rate if the share of low-type consumers is sufficiently high ($\theta < \hat{\theta}_D^C$).

![Graph showing comparison of ISP's profit for one-sided pricing strategies](image)

(a) Maximum heterogeneity: $\alpha_{HA} = \alpha_{BH} = 2.0$. (b) Preference for $A$: $\alpha_{HA} = 2.0, \alpha_{BH} = 1.0$.

Figure 6: Comparison of ISP’s profit for one-sided pricing strategies. (For: $\alpha_L = 1, \beta = d_A = d_B = 1, \gamma = 0$).

More notably, the order of equilibrium profits reverses with respect to plain data caps and zero rating between the two parameter settings. In the case of maximum heterogeneity between both consumer groups, plain data caps weakly dominate zero rating tariffs, whereas the opposite holds true if high-type consumers only have a high valuation for one specific service. Although, the magnitude of the profit difference between the two pricing schemes obviously depends on the distribution of consumers, the relative order is unambiguous for all shares of high-type consumers. Formally, $\theta = 0$ is the unique solution for $\pi^Z_{I} = \pi^C_{I}$ in cases where low-type consumers are not excluded from the market. Therefore, profit functions do not intersect for any $\theta > 0$, unless the ISP reverts to a flat rate offer to high-type consumers in both cases.

Instead, the intra-group heterogeneity, i.e., the degree of heterogeneity in valuations
of high-type consumers, has a crucial impact on the relative profitability of zero rating compared to plain data caps. To see this, let us again fix $\alpha_L$ and $\alpha_{A_H}$ such that consumer groups are sufficiently heterogeneous with respect to their valuation of CSP $A$. Then there is a unique threshold for high-type consumers’ valuation of CSP $B$ $\hat{\alpha}_{H_B} \in [\alpha_L, \alpha_{A_H}]$, where profit functions of zero rating and plain data caps intersect. In consequence, plain data caps are more profitable than zero rating if the intra-group heterogeneity for high-type consumers is relatively small, i.e., $\alpha_{H_B} \in (\hat{\alpha}_{H_B}, \alpha_{A_H}]$. In contrast, zero rating tariffs dominate plain data caps if the intra-group heterogeneity is relatively large, i.e., $\alpha_{H_B} \in [\alpha_L, \hat{\alpha}_{H_B})$. To derive $\hat{\alpha}_{H_B}$, we solve $\pi^Z = \pi^D$ and obtain

$$\hat{\alpha}_{H_B} := \frac{1}{(-\beta d_B + 2 d_A \gamma)} \left( ((\alpha_{H_A} - \alpha_L)d_A - d_B \alpha_L) \beta + 2d_A \gamma \alpha_L \right)$$

- $\sqrt{-\beta (\alpha_L - \alpha_{H_A})^- (\beta d_A^2 - \beta d_B^2 + 2d_A d_B \gamma)}$.}

**Result 4.** Zero rating is profitable for the ISP compared to plain data caps if consumer segments mainly differ in their valuation for specific content, and are relatively homogeneous with respect to other content. In that case, the ISP is able to increase profits by zero rating if and only if $\alpha_{H_B} < \hat{\alpha}_{H_B}$. Zero rating of CSP $B$ allows the ISP to better price discriminate by means of data caps, because these caps can then be designed with respect to the more heterogeneous consumption of CSP $A$.

**Intra-group heterogeneity within the low-value consumer group:** For ease of exposition we have so far assumed heterogeneity within the high-value consumer group, and have assumed that low-value consumers are homogeneous with respect to their preferences for CSP $A$ and CSP $B$. As shown by the previous result, the ISP can profitably use zero rating to make consumer groups more distinct from each other by eliminating the service from the traffic demand pattern that otherwise makes consumer groups more similar to
each other in their average data consumption (see also the results obtained by Inceoglu and Liu, 2017). In our model setup the ISP will therefore zero rate the CSP that is (on average) less desired by consumers. If we would, in contrast assume heterogeneity of low-value consumers and homogeneity of high-value consumers, the zero rating decision of the ISP may change.

![Preference Matrix](image)

Figure 7: Intra-group heterogeneity within the low-value consumer group.

To see this, consider the scenario depicted in Figure 7. Again we assume that valuations within the heterogeneous group (now the low-value consumer segment) are higher for CSP A and lower for CSP B, i.e., \( \alpha_{LA} \geq \alpha_{LB} \). As a result consumers are more similar with respect to CSP A, which is on average the more valuable CSP from the perspective of consumers. To make both consumer groups more distinct with respect to the observed data traffic pattern, the ISP would in this case zero rate the higher-valued CSP A. However, the main insight derived from our zero rating analysis still holds true under this alternative specification of consumer preferences.

**Result 5.** Zero rating may allow for a better separation of consumer segments (i.e., low and high consumer types) by eliminating the service for which preferences are most similar between segments.

- If high-value consumers are heterogeneous and low-value consumers are homogeneous the ISP prefers to zero rate the lower-valued CSP.
If low-value consumers are heterogeneous and high-value consumers are homogeneous the ISP prefers to zero rate the higher-valued CSP.

In both cases the remaining data traffic for the non-zero rated CSP becomes more heterogeneous between consumer groups, which may allow the ISP to better price discriminate.

**CSP profit:** Next, we examine possible effects of zero rating and sponsored data on the content and services market. Figure 8 presents the profit of both CSPs in the case of each ISP pricing scheme. For the denoted parameter setting with $\alpha_{HA} > \alpha_{HB}$, CSP $A$ (see Figure 8a) makes a higher profit than CSP $B$ (see Figure 8b) in the case of a flat rate or plain data caps, because high-value consumers have a preference for the former and the ISP treats traffic of both CSPs equally. However, in the case of (exclusive) zero rating and sponsored data, as described in Section 4, traffic of CSP $B$ is unconstrained, whereas traffic of CSP $A$ must be consumed within the limits of the respective data cap. In the denoted parameter setting, the ISP strictly prefers zero rating over plain data caps and will therefore zero rate CSP $B$ if possible, as shown in Figure B1 in the Appendix.

Remarkably, CSP $B$, the less-valued service in this case, does not only (weakly) increase its profit under zero rating relative to any other pricing scheme, but also obtains higher profits than CSP $A$ if the share of high-value consumers is not too high. Moreover, CSP $A$ may not only be worse off relative to the other CSP, but also obtains (weakly) lower profits than in the case of plain data caps without zero rating. Therefore, zero rating may have important ramifications for CSPs even if they themselves are not zero rated and, most notably, even if they are not in direct competition with CSPs that are zero rated.\(^5\) This finding can be explained by the general notion that under a binding data cap, even independent content and services find themselves in a competitive relationship from the perspective of consumers, as highlighted by Economides and Hermalin (2015). Finally,

\(^5\)Note that we have considered only independent CSPs (i.e., $\gamma = 0$) so far.
from the perspective of other CSPs, it is irrelevant whether CSP $B$ is zero rated for free or by the means of a sponsored data deal. However, those CSPs may still be worse off in settings where sponsored data is allowed, because the ISP is then more likely to exempt traffic of CSP $B$ from the data cap.

**Result 6.** The zero rated CSP weakly prefers zero rating to all other pricing schemes. However, zero rating makes the CSP, that is not zero rated, weakly worse off. This is even the case if CSPs are independent, i.e., for $\gamma = 0$. Moreover, zero rating may distort demand of consumers, such that they consume more of the less-preferred service. If CPSs are advertising-financed, this can lead to higher profits for the inferior CSP.

**Cap sizes and consumer surplus:** The public and regulatory debate on zero rating has often focused on the issue of cap sizes. In particular, opponents have pointed to the threat that under zero rating and sponsored data, cap sizes may be shrinking and therefore
internet users may be worse off even if they could consume some content in an unconstrained manner. The previous result on CSPs’ profit has already provided some indication, that zero rating may indeed have a significant effect on the consumption patterns of consumers. However, as shown below, such metrics are likely to be inconclusive with respect to actual consumer welfare.

Consumer surplus $W_{CS}$ is calculated as the aggregated net utility over all consumers in the market. As all pricing schemes either extract the entire surplus of low-type consumers or exclude them from the market, positive consumer surplus can only stem from high-type consumers. Thus, calculation of consumer surplus reduces to $W_{CS} = \theta(U_H[q_A, q_B] - F_H)$.

The left panel of both Figure 9 and Figure 10 depicts cap sizes of tariffs designed for low-type consumers $C_L$ in the case of zero rating and plain data caps. In both cases, the cap size with zero rating is (weakly) smaller than the cap size with plain data caps. But, as traffic of CSP $B$ does not count against the cap under zero rating, the cap size itself does not convey any clear implications. However, as the data cap is now used exclusively for traffic of CSP $A$, the cap size for zero rating ($C^{ZR}_L$) can be compared to the data traffic for CSP $A$ under the plain data cap ($D^{DC}_{LA}$). In both figures, the data cap size under zero rating is below this traffic benchmark. Thus, under zero rating, the data cap is more constraining than the plain data cap with regard to consumption of CSP $A$ for low-type consumers.

The right panel of both Figure 9 and Figure 10 denotes the consumer surplus for the respective situation that we have just discussed with respect to cap sizes. Although, cap sizes under zero rating are more constraining than plain data caps in both depicted situations, consumer surplus may either decrease (as in Figure 9b) or increase (as in Figure 10b).

**Result 7.** Lower cap sizes under zero rating are not necessarily indicative of lower consumer surplus, even if they constrain consumption of the non-zero rated CSP more than a plain data cap.

In the situation depicted in Figure 9, zero rating would leave consumers unambiguously
Figure 9: Comparison of (consumer welfare-decreasing) zero rating and plain data caps.
(For: $\alpha_{H_A} = 2, \alpha_{B_H} = 1.1, \alpha_L = 1, \beta = d_A = d_B = 1, \gamma = 0, r = 2$).

Figure 10: Comparison of (consumer welfare-increasing) zero rating and plain data caps.
(For: $\alpha_{H_A} = 2, \alpha_{B_H} = 1.6, \alpha_L = 1, \beta = d_A = d_B = 1, \gamma = 0, r = 2$).
better off in aggregate. This is because it is more difficult for the ISP to satisfy the incentive constraint with zero rating and therefore high-type consumers face a lower subscription fee $F_H^{ZR}$ than under plain data caps. Yet, this is also the reason, why the ISP finds it profitable to offer a zero rating tariff over a tariff with plain data caps in the situation depicted in Figure 9, but not in the situation depicted in Figure 10.

**Sponsored data:** To examine the welfare implications of sponsored data we analyze the effects on consumer surplus and total welfare as denoted in Figure 11 for two numerical examples. We calculate total welfare $W_{TS}$ as the sum of consumer surplus and profits of all firms, i.e., the ISP as well as both CSPs. As argued in Section 4.4, the profit of the ISP may increase with sponsored data, however, from a welfare perspective, this is a mere shift of rents from the sponsoring CSP to the ISP. Therefore, welfare effects that differ from zero rating must either stem from changes in consumer surplus or the surplus of CSP $A$, which does not participate in the sponsored data deal. We already know from the analysis of zero rating that the latter will indeed be (weakly) worse off compared to plain data caps.

With respect to consumers, Figure 11b reveals that sponsored data can in fact lead to an increase in consumer surplus in regions where the ISP finds it profitable to make a sponsored data offer to the CSP. This gain in surplus is driven by the fact that zero rating would not be profitable for the ISP in this region, but beneficial to consumers. Because sponsored data has the same implications for consumers’ demand as zero rating, the ability of the ISP to effectively price discriminate between both consumer segments is diminished. This in turn directly translates into a gain in consumer surplus, which carries over to total welfare, as can be seen in Figure 11d.

**Result 8.** Sponsored data, if offered by the ISP (and accepted by the CSP), weakly increases consumer surplus and total welfare.

Although sponsored data increases consumer and total surplus, it distorts consumption
(a) Consumer surplus for $\alpha_{BH} = 1.1$.

(b) Consumer surplus for $\alpha_{BH} = 1.6$.

(c) Total surplus for $\alpha_{BH} = 1.1$.

(d) Total surplus for $\alpha_{BH} = 1.6$.

Figure 11: Welfare analysis for sponsored data. (For: $\alpha_{HA} = 2, \alpha_{L} = 1, \beta = d_{A} = d_{B} = 1, \gamma = 0, r = 2$).
patterns of consumers with respect to CSPs and therefore affects the profits of both CSPs. With regard to dynamic efficiency and fairness concerns, this may represent an objective justifications for prohibiting sponsored data. However, based on the previous result, this may imply both a welfare loss and a consumer surplus loss in cases, where sponsored data would have been implemented instead of plain data caps. That implies that a regulator cannot maintain the same level of consumer surplus and total surplus if sponsored data agreements are prohibited per se, because the ISP would then, in several cases, implement data caps instead of zero rating.

6 Conclusion

This study contributes to the growing literature on zero rating and sponsored data as well as the larger strand on data caps and the strategic use of download limits. In this context, we show how artificial bottlenecks as well as exemptions from those can be used by a monopolistic ISP to maximize profits. In our analysis we distinguish between zero rating, a one-sided pricing mechanism and sponsored data, a two-sided pricing mechanism. We propose a theoretical model that exemplifies how data caps in combination with zero rating can be used to better price discriminate between heterogeneous consumer groups depending on the level of heterogeneity in the population.

In this context, our findings show that varying data cap sizes are generally not an adequate indicator for changes in consumer surplus. Even in the case that data caps are more restrictive under zero rating compared to plain data caps, there are situations where consumers unambiguously benefit from zero rating. More specifically, this is the case if consumer groups are rather heterogeneous in their valuation for zero rated content. Regulators should therefore refrain from using data cap size comparisons (e.g., between different countries) to make judgments about consumer welfare in the case of zero rating.
Moreover, we find that zero rating can increase total welfare, if different consumer segments are not too homogeneous in their preference for zero rated content. In contrast, zero rating may also harm consumers if the share of high-value consumers is relatively low and consumer segments are rather homogeneous in their valuations for zero rated content.

Even more than zero rating, sponsored data has raised numerous net neutrality concerns among regulators. We find that sponsored data, if offered by the ISP, increases consumer surplus and total welfare. However, we also show that sponsored data distorts the level playing field between CSPs in favor of the sponsoring CSP and may cause a loss for CSPs that are not part of the sponsoring deal. This is even the case if CSPs are not in direct competition with each other (i.e., if CSPs are independent). Therefore, in line with previous studies, we argue that data caps introduces competition between CSPs that are not in the same market for content and services. In addition, sponsored data may create asymmetric rationing schemes of Internet content, where some content providers are subject to such download limits, whereas others are not. However, these effects on the CSP market are not limited to sponsored data, but also extend to zero rating. Therefore, we challenge the consensus in regulatory practice that sponsored data should generally be met with more regulatory scrutiny than zero rating.

In closing, we wish to point out some limitations of our analysis that call for future research. First, this study, in line with the related literature, focuses on zero rating and sponsored data in the context of a monopoly ISP. However, especially with regard to zero rating, competitive pressure between ISPs is often cited as a driving force behind marketing decisions to exempt specific services or entire services categories from a user’s mobile data allowance. Thus, on the one hand, competition between ISPs may lead to more zero rating offers. On the other hand, zero rating may become less likely, as competition constrains an ISP’s market power, which is a necessary prerequisite for extracting rents through price discrimination. Thus, future work may examine the implications of these
two countervailing effects.

Second, we consider advertising-financed content and services providers, which account for an important share of the Internet ecosystem. However, zero rating offers have also been popular in the case of data-intensive video services (see, e.g., Netflix), which are often financed through a monthly subscription model. Jullien and Sand-Zantman (2018) examine paid content in the context of sponsored data and argue that in this case the price structure is irrelevant for the market outcome. Thus, they conclude that there is no scope for any sponsored data option. Nonetheless, the practical relevance of zero rating for paid content may warrant additional analyses.

Third, our study abstracts from capacity constraints, investments and network congestion. These issues may be relevant in the case of zero rating and sponsored data, as specific services and content (categories) can now be consumed unconstrained by Internet users. Therefore, the likelihood of congestion and also the necessity of capacity investments may increase. Moreover, additional profits from these pricing schemes may lead to more capacity investments (see, e.g., Inceoglu and Liu, 2017). In contrast, ISPs mitigate data traffic expansion in practice, as they usually implement zero rating schemes together with throttling techniques that, e.g., reduce the quality of video streams and thus reduce peak traffic loads. As our model offers a microfoundation for data cap tariffs in the context of zero rating and allows to explicitly consider the traffic intensity of specific services, this study may serve as a basis for future research in this regard.

References


Federal Communications Commission (2017). Wireless telecommunications bureau report: Policy review of mobile broadband operators’ sponsored data offerings for zero-rated con-


Appendix

A Zero rating

A.1 Zero rating of CSP B

Assuming that the ISP serves the entire market, i.e., both consumer segments, the ISP maximizes its profit by setting cap sizes $C_{H}^{ZR}$ and $C_{L}^{ZR}$. As shown in Section 4.3, the ISP then sets a positive data cap for low-type consumers for $\theta > \theta_{0}^{ZR}$ and a cap size of zero if $\theta_{0}^{ZR} \leq \theta$.

Accordingly, the optimal fixed fee of the tariff designed for low-value consumers is as follows:

$$F_{L}^{ZR} = \begin{cases} F_{L}^{ZR,1} & \text{if } 0 \leq \theta < \theta_{0}^{ZR} \\ \frac{\alpha_{L}}{\theta_{3}} & \text{if } \theta_{0}^{ZR} \leq \theta < 1, \end{cases}$$
with
\[
F^{ZR,1}_L = \frac{1}{2(-\beta^2 + \gamma^2)(-1 + \theta)^2}((\alpha_{HA}^2 - 2\alpha_{HA}\alpha_L - \alpha_L^2)\beta^2
+ 2((\alpha_{HA} + \alpha_{HB})\alpha_L - \alpha_{HA}\alpha_{HB})\gamma\beta + \gamma^2(\alpha_L - \alpha_{HB})^2)\theta^2
- 4\beta\alpha_L^2(\gamma - \beta) + 2\beta\alpha_L^2(\gamma - \beta).
\]

The fixed fee of the tariff designed for high-value consumers, subject to the assumption that it is optimal to let all consumer segments participate in the market, is given by
\[
F^{ZR}_H = \begin{cases} 
F^{ZR,1}_H & \text{if } 0 \leq \theta < \tilde{\theta}^{ZR} \\
\frac{((-\alpha_{HA}^2 - \alpha_L^2)\beta^2 + 2\gamma\beta\alpha_{HA}\alpha_{HB} + \gamma^2(-\alpha_{HB}^2 + \alpha_L^2))}{(-2\beta^4 + 2\beta\gamma^2)} & \text{if } \tilde{\theta}^{ZR} \leq \theta < 1,
\end{cases}
\]

with
\[
F^{ZR,1}_H = \frac{1}{((2(-1 + \theta))(-\beta^2 + \gamma^2)^2\beta)}((\gamma - \beta)((-\theta + 3)\alpha_L^2 - 2\alpha_{HA}((\theta + 1)\alpha_L
+ \alpha_{HA}^2(\theta + 1)))\beta^2 + 2(-2\alpha_L^2 + (\alpha_{HA} + \alpha_{HB})(\theta + 1)\alpha_L - \alpha_{HA}\alpha_{HB}((\theta + 1)))\gamma\beta
+ \gamma^2(\alpha_L - \alpha_{HB}^2(\theta + 1))(\beta + \gamma)).
\]

However, the ISP may further increase profits if it excludes low-type consumers in the market and extracts the entire surplus from high-type consumers that receive a quasi flat rate, i.e., a data cap that allows for their unconstrained consumption of CSP $A$ and zero rating of CSP $B$. This is because under zero rating high-type consumers may still have an incentive to disguise as low-type consumers for a cap size of zero in order to benefit from the zero rating of service $B$ at a lower fixed fee. In consequence, the ISP cannot charge high-type consumers the fixed fee that would extract the entire surplus as long as low-type consumers still participate in the market.
For the first case, $0 \leq \theta < \hat{\theta}^{ZR}$, the ISP will exclude low-types from the market if and only if $\theta > \theta_H^{ZR1}$. For the second case, $\hat{\theta}^{ZR} \leq \theta \leq 1$, the ISP will exclude low-types from the market if and only if $\theta > \theta_H^{ZR2}$. In both cases, the ISP, under exclusion, would set $F_L^{ZR} > U_L[0, q_B^{ZR}(0)]$ and $F_H^{ZR} = \frac{\alpha_L^2}{\beta + \gamma}$ to maximize profits.

Taking into account the ISP’s option to exclude low-type consumers from the market, optimal profits under zero rating are then given by

$$
\hat{\theta}_H^{ZR1} := \frac{1}{-2\beta^2\alpha_A^2 - 2\beta^2\alpha_B^2 + 4\gamma\beta\alpha_A\alpha_B\alpha_L + 4\gamma^2\alpha_A^2\alpha_B^2 + 2\gamma\alpha_L(\alpha_L - \alpha_B^2)}(\theta^2 + \gamma^2)\alpha_L^2 - 2(\gamma - \beta)(\beta\alpha_A + \gamma\alpha_B)\alpha_L
$$

$$+ (-\beta^2 + \gamma^2)\alpha_B^2 + ((-\alpha_B^4 + (-4\alpha_A\alpha_L + 6\alpha_L^2)\alpha_B^2 - \alpha_L^4 - 4\alpha_A^3\alpha_L + 4\alpha_B^2\alpha_L^2)\beta^2
+ 4\gamma\alpha_L(\alpha_L - \alpha_B)\alpha_L^2(\alpha_L + \alpha_B)\beta + \gamma^2(\alpha_L - \alpha_B)^2(\gamma - \beta)(\beta + \gamma))^{\frac{1}{2}},
$$

$$
\hat{\theta}_H^{ZR2} := \frac{\alpha_L^2}{\alpha_B^2}.
$$

$$
\pi_I^{ZR} = \begin{cases} 
\pi_I^{ZR,1} & \text{if } \theta \leq \theta_H^{ZR1} \\
(\theta(\beta\alpha_A^2 + \beta\alpha_B^2 - 2\gamma\alpha_A\alpha_B)) & \text{if } 0 \leq \theta < \hat{\theta}^{ZR} \\
(\pi_I^{ZR,2}) & \text{if } \theta \leq \theta_H^{ZR2} \leq \theta < 1,
\end{cases}
$$

with

$$
\pi_I^{ZR,1} := \frac{1}{(2\beta(\theta - 1)(-\beta^2 + \gamma^2))}(((\alpha_A^2 - 2\alpha_A\alpha_L - \alpha_L^2)\beta^2 + 2\gamma(\alpha_A^2 - \alpha_B\alpha_B)\beta + \gamma^2(\alpha_L - \alpha_B)^2)\theta - 2\beta\alpha_L^2(\gamma - \beta)),$$

$$
\pi_I^{ZR,2} := \frac{((-\theta\alpha_A^2 - \alpha_L^2)\beta^2 + 2\beta\gamma\theta\alpha_A\alpha_B^2 - \gamma^2(\theta\alpha_B^2 - \alpha_L^2))}{(-2\beta^3 + 2\beta\gamma^2)}.
$$
A.2 Zero rating of CSP A

Assuming that the ISP serves the entire market, i.e., low-type consumers are not excluded by means of a high subscription fee, the optimal profit of the ISP is given by

$$\pi_{IA}^{ZRA} = \begin{cases} 
\pi_{IA}^{ZRA,1} & \text{if } 0 \leq \theta < \tilde{\theta}^{ZRA} \\
\pi_{IA}^{ZRA,2} & \text{if } \tilde{\theta}^{ZRA} \leq \theta < 1,
\end{cases}$$

with

$$\pi_{IA}^{ZRA,1} := \frac{1}{(2\beta(\theta - 1)(-\beta^2 + \gamma^2))} \left((\alpha_H^2 - 2\alpha_H^2)\alpha_L - \alpha_L^2\right)^2 \beta^2 + 2\gamma((\alpha_H + \alpha_H^2)\alpha_L - \alpha_H^2)\beta + \gamma^2(\alpha_L - \alpha_H^2)\theta - 2\beta\alpha_L^2(\gamma - \beta)), $$

$$\pi_{IA}^{ZRA,2} := \frac{(\alpha_H^2 - \alpha_L^2)^2 + 2\beta\gamma\theta\alpha_H - \gamma^2(\theta\alpha_H^2 - \alpha_L^2))}{(-2\beta^2 + 2\beta\gamma^2)},$$

$$\tilde{\theta}^{ZRA} := \frac{\alpha_L(\beta - \gamma)}{(\beta\alpha_H - \gamma\alpha_H)}. $$
B Supplementary figures

Figure B1: ISP’s profit for one-sided pricing strategies. (For: \( \alpha_{HA} = 2, \alpha_{BH} = 1.1, \alpha_L = 1, \beta = d_A = d_B = 1, \gamma = 0, r = 2, s = 0.1 \)).