Internet Regulation, Two-Sided Pricing, and Sponsored Data *

Bruno Jullien[†] Wilfried Sand-Zantman[‡]

March 2017

Abstract

We consider a network that intermediates traffic between the consumers and providers of free content. We analyze the implications of offering sponsored data plans that allow content providers to pay for traffic on behalf of their consumers. Sponsored data boosts consumption of high-value content, but the network may charge higher prices to consumers for non-sponsored content. The welfare effects of allowing sponsored data depend on the proportion of content targeted and the value of such content. Our analysis is conducted under two-sided prices and under one-sided pricing (only consumers pay), and it is extended to the case of network competition.

1 Introduction

The pricing of traffic on the Internet is the subject of many controversies due to contrasting views on how Internet service providers (ISPs), which manage the physical network, should treat various types of content and on their relationship with content providers. In this paper, we discuss the role of traffic management methods that allow content providers to pay for the

*This paper circulated previously under the title “Congestion Pricing and Net Neutrality". We thank two referees and Martin Peitz for excellent reviews. We are grateful to Dominik Grafenhofer, Byung-Cheol Kim, Jan Krämer, Martin Peitz, Mike Riordan, David Salant, Florian Schuett, Yossi Spiegel and Tommaso Valletti for helpful discussions and comments, as well as participants at CRESSE Conference, CREST-LEI seminar, The Future of Internet Conference (MaCCI, Mannheim), ICT Conference (ParisTech), EARIE, the 2nd London IO Day and Pontifica Univerdad de Chile seminar. We gratefully acknowledge Orange, in particular Marc Lebourges, for its intellectual and financial support under the IDEI/Orange convention. Wilfried Sand-Zantman also acknowledges support from the Agence Nationale de la Recherche, grant ANR-12-BSH1-0009.

†Toulouse School of Economics (CNRS and IDEI). E-mail: bruno.jullien@tse-fr.eu

‡Toulouse School of Economics (GREMAQ and IDEI) and ESSEC Business School. Manufacture des Tabacs, 21 allées de Brienne, 31000 Toulouse. E-mail: wsandz@tse-fr.eu

data used by their customers. Examples of such methods include agreements such as AT&T's *Sponsored Data* plan, Verizon FreeBee Data, and other *zero-rating* programs, whereby the data generated by the consumers of a provider are not counted in the subscriber's own monthly data limits.[1] Zero-rating plans are offered not only in developed countries but also in developing countries, through partnerships between large application developers and mobile operators, such as the programs Free Basics by Facebook or Google Free Zone. These types of agreements have raised concerns among net neutrality advocates and are the object of an intense debate between Internet actors and regulators.[2] The issue is whether discriminatory agreements may be justified by efficiency considerations.

To see how this problem differs from standard discriminatory pricing problems, let us begin by adopting a global perspective on the optimal pricing strategy for consumer content. The Internet can be seen as a three-party business whereby content providers and consumers use the network to trade. This trade creates some costs (mostly for the network) and some benefits both for consumers and content providers, either directly, such as when consumers pay for usage, or indirectly, such as when the presence of consumers generates ad revenues. When the presence of these consumers results in either costs or benefits, it is natural to suppose that content providers and the network could adjust their respective prices to deter or foster consumer usage. However, there are two major factors that impede such adjustment. First, many websites are free and thus cannot use prices to induce the behavior desired from consumers. Second, websites are prone to differ in terms not only of the cost they impose on the network but also the benefits they create. The combination of a "missing price" for content and heterogeneity implies that benefits and costs will not be internalized properly by consumers or the network, which compromises the efficient use of network capacity. In this work, we derive the network's optimal pricing strategy when websites differ in the social value they generate and cannot directly affect consumer behavior via price. We explore how network tariffs that target both consumers and content providers can be designed to alleviate the misallocation problem of network capacity and promote efficient network use.

Toward this end, we model a network that intermediates the traffic between content providers and consumers. Content providers receive a benefit proportional to traffic, such as advertising revenue or direct utility for the producer. This benefit is heterogenous, with high-benefit and low-benefit content providers, but is private information of the content providers. Furthermore, the total benefits depend also on the usage level chosen by consumers. Because

[1] See *https://developer.att.com/sponsored-data* for AT&T's plan, and *https://www.5screensmedia.com/enterprise* or *https://www.sandvine.com/solutions/subscriber-services/sponsored-data.html* for platforms intermediating such services.

[2] Regulators have declared zero rating to be anti-competitive in Canada, Chile, India, Norway, the Netherlands, and Slovenia. See the OECD's *Digital Economy Outlook* for 2015.

consumer usage increases the network's total cost, a price can be charged to one or both of the parties involved in traffic generation. When only consumers are charged, we refer to one-sided pricing, whereas when both consumers and content providers are charged, we refer to two-sided pricing. We focus on the two-sided pricing case but demonstrate in the extension that our insights apply also to the one-sided case. To match actual practice, we assume that consumers pay a tariff that depends on the data allowance they choose while the content producers are charged a non-discriminatory linear price for the traffic they generate.

We also allow the network to propose a "sponsored data" plan to the content providers whereby the consumer traffic to their website is removed from the data allowance. By relaxing the constraint on traffic, this influences the consumption choice and raises traffic and, therefore, the website's revenues. Given the network's option, each content provider must trade off the volume of consumption against the cost of traffic. The high-benefit content providers will sponsor consumption to generate higher advertising revenues, while low-benefit providers will prefer to reduce their costs by allowing consumers to pay for traffic. Such a sponsored plan thus enables the network to discriminate among various content types. These practices emerge naturally as a correction for allocative inefficiencies arising from the absence of some prices (here, the price of content).

With paid content, the price structure between users and providers is irrelevant. The same efficient outcome can be implemented under one-sided and two-sided pricing. Furthermore, there is no scope for any sponsored data option.

With free content, absent a sponsored data program for the content providers, the network may decide to exclude the low-value providers to extract more from the high-value providers. Allowing a sponsored data program alleviates this standard monopoly trade-off. Providers have the option to increase consumption by paying for traffic, which allows the network to differentiate consumption by content type. Despite improving efficiency compared with uniform pricing, the menu of prices results in socially suboptimal consumption levels. Indeed, the network reduces consumption levels under the base tariff to raise the relative value and thus the price of sponsored data.

In this paper, we examine the effects of a regulation banning sponsored data — that is, of forcing the network to offer a single pair of tariffs with one part to be paid by consumers and the other by content providers. Such uniform pricing prevents the network from adjusting data allowances – or consumption – to the type of content. Since rents cannot be extracted from the most valuable content without excluding the low-benefit content, it follows that there will be more exclusion under uniform pricing than in the case of sponsored data. Absent exclusion, the welfare consequences of the ban are ambiguous due to the opposite effects on the two types of content. We show that a ban on sponsored data reduces welfare if the proportion of low-benefit content is large, but it raises welfare if the value of low-benefit content is not too low.

We extend the analysis in several directions. First, we consider the case of one-sided pricing. Sponsored data generates higher incremental value for adopters of the plan than under two-sided pricing, whereas the consumer price for data increases more. The welfare analysis is similar, but there cannot be exclusion of any content in this case.

Second, we discuss the case of elastic demand and competing networks. We show that competing networks will still choose to propose the sponsored data program derived in the benchmark case to maximize their profits. We also show that elastic consumer participation argues for less regulation because the network will pass on to consumers a larger share of the gains obtained from content providers.

Third, we discuss a more general model that accommodates content providers differing in terms of demand, advertising benefits and traffic intensity. We show that the analysis extends to this case and that what matters for the adoption of sponsored data is the content providers' benefit per unit of traffic generated.

Finally we extend the analysis to the more general case in which paid content and free content coexist.

Our work is first related to the literature on two-sided markets (see Caillaud and Jullien, 2003; Rochet and Tirole, 2003, 2006; Armstrong, 2006), in that we aim to characterize the optimal pricing for each side of the market, content providers and consumers. We combine the participation model of Armstrong (2006) and the usage model of Rochet and Tirole (2003). In our basic model, the total number of agents on the consumer side is fixed and their consumption is affected only by the price and number — or more precisely the types — of content providers in the market. On the content-provider side, profit depends not only on the network charges providers must pay but also on the number of consumers and the price that they are charged by the network. The number of content providers can vary, in part because low-benefit providers may be priced out of the market by network charges.

Some contributions in the field of telecommunications have studied the sender–receiver pricing structure. This literature emphasizes the importance of "call externalities" and, hence, the social benefits associated with using positive receiver prices (Jeon, Laffont and Tirole, 2004). Hermalin and Katz (2004) develop a related idea while focusing on how best to address uncertainty over the private value of exchanging messages, and the game — namely the choice to call or to wait for a call — induced by the tariff structure. In our paper, the structure of communication is different because it is the receiver (i.e., the consumer) who always initiates the communication. Another difference is that, in our setup, only the sender learns the true benefit of this communication.

The literature on the Internet price regulation has been driven by the debate over net neutrality and the optimal way to price content providers and consumers (see, e.g., Economides and Hermalin, 2012, and Greenstein, Peitz and Valletti, 2016). One point emerging from two-sided market models is that, although laissez-faire can be shown to result in inefficient pricing, the precise nature of an intervention that would foster efficiency is unclear (Economides and Tåg, 2012). While neglecting the investment question (on this point, see, e.g., Choi and Kim, 2010; Hermalin and Katz, 2009), we direct our attention to the efficient management of current resources when the benefit of consumption is uncertain. Several recent contributions discuss the screening of traffic-sensitive content by means of prices and differentiated quality layers, a key aspect of the net neutrality debate (Krämer and Wiewiorra, 2012; Reggiani and Valletti, 2016; Choi, Jeon and Kim, 2015; Bourreau, Kourandi and Valletti, 2015). Peitz and Schuett (2015) analyze moral hazard in traffic generation using a model that incorporates congestion externalities.

Our work departs from these papers by considering consumption usage and the allocative role of consumer prices. A specific contribution is to show when a sponsored data plan permits screening among traffic-sensitive content types and enhanced efficiency.

The remainder of the paper proceeds as follows. After describing the model and discussing paid content in Section 2, in Section 3 we analyze the case of free content and two-sided pricing, where we discuss the consequences of a ban on sponsored data. In Section 4, we explore the extensions mentioned above. Section 5 concludes.

2 Model

2.1 Framework

We analyze the tariff charged for traffic by a network (an ISP, in the case of the Internet) to two sides of the market: a unit mass of consumers and a unit mass of content providers. In practice, some content is delivered freely while other content is paid for. To focus on our paper's novel aspects, we simplify the analysis by assuming that all content is free, but we will present briefly the case of paid content for benchmarking. We assume for conciseness that content is non-rival such that consumers visit every content provider. The expected demand for each content when consumers face a price p per unit of content is $q = D(p)$. The representative consumer's utility function $u(q)$ is strictly concave on $[0, \bar{q}]$ with $u'(0) = \bar{p} > 0$ and $u'(\bar{q}) = 0$. Thus, demand $D(\cdot)$ is decreasing, maximal consumption $D(0) = \bar{q}$ is positive, and demand vanishes at price $\bar{p}$.

While we model the demand as continuous and homogenous, our setup can be interpreted as well as modelling subscription based content services. For this interpretation, suppose that all type-t contents are ex-ante homogenous (before joining the network) but that ex-post, when choosing the consumption of each content, a consumer has a unit demand with $i.i.d.$ random reservation price $\tilde{u}$. Then the price p is a subscription price while $q = D(p)$ is the probability that $\tilde{u} \geq p$ and $u(q)$ is the ex-ante expected gross surplus when the consumer anticipates a probability q of subscribing to the content service.

Any transaction between a content provider and a consumer creates costs and benefits in addition to the utility that consumers derive from usage. More precisely, each unit of content generates traffic θ, which is also the cost born by the network. Hence, the consumption of q units of content costs the network an amount θq. This cost is either direct or a function related to congestion. One interpretation is that θ reflects the network's costs of expending resources to maintain service quality. We assume that consumption is positive if priced at its true marginal cost (i.e., $\theta < \bar{p}$). Furthermore, each unit of consumption generates a benefit $a > 0$ for the content providers, net of the cost (if any) of distributing the content. This benefit is diverse, including the advertising revenue[3] and other gains for the content provider such as private benefits for blogs and not-for-profit organizations, the value of consumer data and the leverage of the customer base on the capital market. Content providers are heterogenous and can be of either low-benefit type ℓ or high-benefit type h; these types are characterized by respective benefits a_ℓ and a_h where $a_\ell < a_h$.[4] We assume that any given content is type h with probability λ, and thus, content is of type ℓ with probability $1 - \lambda$. We will refer to content providers of type h as high benefit (in short HB) and to content providers of type ℓ as low benefit (in short LB). We will use $b_\ell = a_\ell/\theta$ and $b_h = a_h/\theta$ to denote the respective ratios of benefits to costs. To focus on the most interesting case, we assume that one type of content would not be profitable if its provider had to pay the full cost of traffic:

$$b_\ell < 1 < b_h.$$

This assumption captures the idea that only some content providers can afford the network's cost. As a consequence, some content providers will not operate unless consumers pay for part of the traffic cost.[5] The network cannot distinguish between the different types; in other words, it is not allowed to practice explicit discrimination.

[3] Advertising revenue increases with consumption if the consumer time devoted to a page is increasing in consumption or if advertising is tied to consumption in some other way.

[4] In the extension section, we will allow consumer utility U and traffic θq to also depend on the website type.

[5] The key assumption is $b_\ell < 1$. Otherwise (as we shall demonstrate), total welfare would be maximized by the imposition of a content price equal to 1.

The network proposes to consumers a choice of data allowance T,[6] that is, a maximal level of traffic consumption, at a price P. Without loss of generality, we assume that P is a two-part tariff of the type $P(T) = F + rT$ where F is base subscription and $r \geq 0$ is the implicit price of data in the tariff.[7]

We assume that the network is connected to competitive backbones and charges a non-discriminatory linear price $s \geq 0$ for the traffic generated by a website. Competition ensures that the price s is charged to each website, such that each content provider's cost of servicing a quantity q of content to a consumer is $s\theta q$. However, below, we will allow the network to offer content providers an option for sponsored data that removes the consumers' traffic from the contractual data allowance, in exchange for a payment by the content provider.

Note that as the sponsored consumption is free of charge, it is predictable, equal to the demand $\bar{q}$ at a zero price for the data allowance. Thus, the sponsored traffic is $\theta\bar{q}$. For convenience, we denote by $\underline{s}$ the price paid by the website per unit of traffic. Thus, the actual payment for sponsored data is $\bar{s}\theta\bar{q}$.

We consider the following timing:

1. The network proposes the prices F, r, s and $\bar{s}$ if the sponsored data program is allowed.

2. Each content provider decides whether to be active and may also choose the sponsored data option, if available.

3. Each consumer decides whether to subscribe, selects a data allowance and then determines how much to consume of each provider's content.

4. Traffic is observed, and payments are made to the network.

For the main part of the paper, we focus on the case of a monopoly network with inelastic participation by consumers. Let us denote by CS the gross consumer surplus from usage. It is given by $\text{CS} = \lambda u(q_h) + (1-\lambda) u(q_\ell)$, where q_h and q_ℓ represent the consumption of HB and LB content, respectively. Because consumers are *ex ante* identical, the network can extract the full surplus by setting the base subscription F such that the total payment P is equal to the gross consumer surplus: $P = \text{CS}$. Accordingly, the network's objective fully internalizes the consumer surplus. Let Π denote the difference between the network's

[6] For an analysis of data caps by ISPs and mobile operators, see Economides and Hermalin (2015).

[7] As in our model consumers are homogenous and anticipate perfectly their traffic, we could allow for a non-linear tariff $P(T)$ without altering the analysis. Then, we would have the implicit marginal price of data $r = P'(T)$ at the equilibrium quantity.

revenues derived from content providers and the total cost of supporting the traffic, i.e., $\Pi = \lambda (s_h\theta - \theta) q_h + (1-\lambda)(s_\ell\theta - \theta) q_\ell$, where s_h and s_ℓ are the unit price paid by each type of content provider. We define the *network value* as $V = \mathrm{CS} + \Pi$ and assume that the network maximizes this network value.[8]

As a first benchmark, let us consider the socially optimal allocation with full information on the content type. This scenario corresponds to the case of a regulated network maximizing social welfare and perfectly discriminating between content types. Content with benefit $a = b\theta$ generates social welfare of $u(q) + (b\theta - \theta)q$. If we ignore the feasibility constraints, then social welfare is maximal at $u'(q) = 1-b$. As consumption cannot exceed $\bar{q}$, consumption is set at this maximal level if the benefit b is larger than 1.

Lemma 1 *The socially optimal consumption q_t^{FB} for any type-t content, $t = \ell, h$, is obtained at $u'\left(q_t^{\mathrm{FB}}\right) = \max\{1-b_t, 0\}\theta$.*

When the content price for a type-t content is $s = b_t$, the content provider receives zero surplus. Hence, for $s_t = b_t$, the network value for each type-t content $V = u(q_t) - \theta q_t + b_t\theta q_t$ is equal to the total welfare. This implies that a network maximizing V implements the social optimum under perfect price discrimination with respect to content.

2.2 Paid content

We now investigate the network's choice of tariffs when content providers charge positive prices to consumers. For this, we assume that each content provider is a monopoly, but the argument extends easily to the case of competition.

Let us consider a tariff (F, r) and prices $p_t(j)$ for each piece of content, where $t = \ell, h$, is the type and j is an index of the content provider. Facing the tariff and these prices, a consumer chooses consumption and a data allowance that maximize

$$\lambda \int_{j=0}^{1} \left(u\left(q_h(j)\right) - p_h(j)\, q_h(j)\right) dj + (1-\lambda) \int_{j=0}^{1} \left(u\left(q_\ell(j)\right) - p_\ell(j)\, q_\ell(j)\right) dj - rT - F.$$

Given that, the optimal data allowance is

$$T = \lambda \int_{j=0}^{1} \theta q_h(j)\, dj + (1-\lambda) \int_{j=0}^{1} \theta q_\ell(j)\, dj,$$

which yields demand $q_t(j) = D\left(p_t(j) + r\theta\right)$ for the content offered by provider j of type t.

[8] We will show that network behavior leads to V also being maximized with competition between networks or elastic demand, although V does not coincide with total network profits in these cases.

The profit of content provider j is then $D\left(p_t\left(j\right)+r\theta\right)\left(p_t\left(j\right)-\left(s-b\right)\theta\right)$. Let us define $p^m\left(c\right)=\arg\max_p D\left(p\right)\left(p-c\right)$ and $q^m\left(c\right)$ as the monopoly price and quantity for a cost c and demand function D. Then the content's price and consumption are given by

$$\begin{aligned} p_t &= p^m\left(\left(r+s-b_t\right)\theta\right)-r\theta \\ q_t &= q^m\left(\left(r+s-b_t\right)\theta\right) \end{aligned} \tag{1}$$

and the total traffic is $T=\lambda\theta q_h+\left(1-\lambda\right)\theta q_\ell$.

As mentioned above, the network can then set the base subscription F to bind the reservation utility, which is such that

$$F+rT=\lambda\left(u\left(q_h\right)-p_hq_h\right)+\left(1-\lambda\right)\left(u\left(q_\ell\right)-p_\ell q_\ell\right)$$

The network can thus capture the full consumer surplus, and the profit is

$$V=\lambda\left(u\left(q_h\right)-p_hq_h\right)+\left(1-\lambda\right)\left(u\left(q_\ell\right)-p_\ell q_\ell\right)+\left(s\theta-\theta\right)\left(\lambda q_h+\left(1-\lambda\right)q_\ell\right),$$

where the price and quantities are given by (1).

It is easy to see that a standard tax-neutrality result applies and that only the total price $r+s$ matters.

Proposition 1 *With paid content, the consumption levels* (q_h, q_ℓ) *and the network value* V *depend only on the total price of data,* $w=r+s$*. As a consequence, the network induces the same allocation under one-sided pricing* $(s=0)$ *and two-sided pricing* $(s,r>0)$*.*

Proof. The network solves

$$\begin{aligned} & \max_{s,r}\lambda\left(u\left(q_h\right)-p_hq_h\right)+\left(1-\lambda\right)\left(u\left(q_\ell\right)-p_\ell q_\ell\right)+\left(s-1\right)\theta\left(\lambda q_h+\left(1-\lambda\right)q_\ell\right) \\ = & \max_{s,r}\lambda\left(u\left(q_h\right)-p^m\left(\left(r+s-b_h\right)\theta\right)q_h\right)+\left(1-\lambda\right)\left(u\left(q_\ell\right)-p^m\left(\left(r+s-b_\ell\right)\theta\right)q_\ell\right) \\ & +\left(r+s-1\right)\theta\left(\lambda q_h+\left(1-\lambda\right)q_\ell\right) \end{aligned}$$

which depends only on the total price $s+r$ through the effect on prices and quantity $q_t=q^m\left(\left(r+s-b_t\right)\theta\right)$. Thus, we can without loss of generality set $s=0$. ■

Thus, when the content is paid, the neutrality result implies that it is irrelevant which side is charged for traffic. Setting $r=0$ would amount to a standard supply contract whereby the content provider buys the traffic from the network and resells its content to the buyer. The difference with a standard vertical chain is that because the network can charge consumers for access, it internalizes the effect of prices on consumers. Thus, the issue of double-marginalization is less severe than within a standard vertical chain. The total price $r+s$ may be above or below cost depending on demand. More precisely, $r+s$ is the solution of the first-order condition:

$$-\lambda\theta\frac{\partial p_h}{\partial s}q_h-(1-\lambda)\,\theta\frac{\partial p_\ell}{\partial s}q_\ell+(s-1)\,\theta\left(\lambda\frac{\partial q_h}{\partial s}+(1-\lambda)\,\frac{\partial q_\ell}{\partial s}\right)+\theta\left(\lambda q_h+(1-\lambda)\,q_\ell\right)=0$$

and it is above cost $(r+s>1)$ if

$$\lambda\left(\frac{\partial p_h}{\partial s}-1\right)q_h+(1-\lambda)\left(\frac{\partial p_\ell}{\partial s}-1\right)q_\ell<0.$$

Thus, the price charged for data traffic is above cost if content's retail price p is not too sensitive to the cost of delivering the content, in particular if the pass-through rate $\frac{\partial p_t}{\partial s}$ is less than 1 for both types of content.

As only the total cost of traffic matters for consumers and content providers, there is no scope for a sponsored data option. Indeed, faced with tariff (F,r,s) and a price $\bar{s}$ for sponsored data, all content providers choose the option that yields the lowest total price; thus, they sponsor the data if $\bar{s}<r+s$ irrespective of their type.

3 Free content: two-sided pricing

We now investigate the network's choice of tariffs when the content is free. In the first part, we focus on a single tariff by assuming that the network sets a uniform price s for all content providers. Then, we will consider the impact of proposing a sponsored data option, whereby a content provider pays for consumption on behalf of its consumers.

3.1 Uniform pricing

Under uniform pricing, a content provider's only decision is whether to participate and — conditional on participation — all face the same price. Because content providers do not charge consumers for the good or service they offer, their profits can only be generated via the benefit a. The price s charged by the network to them cannot be passed through to consumers. Therefore, given a price s, a content provider of type t stays in the market if it anticipates a nonnegative profit — that is, if $s\leq b_t$ for $t=\ell,h$. In particular, if s lies between b_ℓ and b_h, then only the HB content providers participate in the market.

Let us denote by $M\in\{\lambda,1\}$ the mass of active content providers. Facing a tariff

$P(T) = F + rT$ for the data allowance and a zero price for content, the consumer will consume the same quantity of each content $q_h = q_\ell = q = T/M\theta$. Knowing M and θ, choosing a data allowance is equivalent to choosing the consumption q of each content. The utility is thus $Mu(q) - rM\theta q - F$. The demand for the data allowance is then given by $u'(q) = r\theta$, and thus, $q = D(r\theta)$ and $T = M\theta D(r\theta)$.

The network can then charge $F = M(u(q) - r\theta q)$ for $q = D(r\theta)$. It obtains a profit equal to the joint surplus with consumers

$$V = M \times [u(q) + (s-1)\theta q] \quad \text{with} \quad q = D(r\theta).$$

The term in brackets captures the incentives to maximize the per-content joint surplus of the network and consumers for a given value of s. The net data cost per unit of content is $(1-s)\theta$, and internal efficiency is achieved by setting — whenever feasible — a marginal price of the data allowance equal to this cost. Because the participation of content providers is independent of the price charged to consumers, the network chooses

$$r = \max\{1-s, 0\}. \tag{2}$$

Given (2), the choice of the network reduces to choosing the price s that content providers will be charged. Observe that for a given participation level M, the network value increases with the price s. Therefore, the network chooses s by comparing two possible prices for content: the maximal price $s = b_\ell$ that maintains full participation with $r = 1 - b_\ell$ and the maximal price $s = b_h$ that preserves the participation of only HB content providers with $r = 0$ (since $1 - b_h < 0$). The consumption levels in these two cases are then $q_\ell^u = q_\ell^{\mathrm{FB}}$ and $q_h^u = \bar{q}$, which respectively yield the following network values:

$$\begin{aligned} V_\ell^u &= u\left(q_\ell^{FB}\right) + (b_\ell - 1)\,\theta q_\ell^{FB} \quad \text{when } s = b_\ell \\ V_h^u &= \lambda\left[u(\bar{q}) + (b_h - 1)\,\theta\bar{q}\right] \quad \text{when } s = b_h. \end{aligned}$$

When both types of content providers participate, the network must leave a higher rent to the HB content providers because it cannot selectively reduce q_ℓ. Hence, there is exclusion of content when the proportion of HB content is large.

Proposition 2 *Under uniform pricing, the network excludes the LB content ($s = b_h$) if and only if $\lambda > \lambda^u$. The threshold λ^u is decreasing in b_h and increasing in b_ℓ.*

Proof. See Appendix. ■

The comparative statics underlying the trade-off can be easily analyzed in terms of the relative efficiency of content types. The network value under exclusion increases with b_h and is independent of b_ℓ. Conversely, the network value when all content providers participate increases with b_ℓ and is independent of b_h. In sum, (i) exclusion occurs if b_h is large enough and/or b_ℓ is small enough, and (ii) the network does *not* exclude the LB content if both b_h and b_ℓ are close to 1. Note that, when $b_l = 0$, λ^u is positive and for $\lambda < \lambda^u$, the network chooses to set a price equal to zero to content providers.

3.2 Sponsored data

With uniform pricing, the network faces the standard monopoly trade-off between capturing the rent of HB content providers (with high s) and avoiding the exclusion of LB content providers (with low s). One way to alleviate this trade-off consists in allowing the network to propose tariffs that are more complex. Therefore, we now consider the possibility of the network achieving second-degree price discrimination between the two types of content providers by offering a sponsored data option.

Sponsored data option: The network gives content providers the option to remove their traffic from consumers' data allowance and instead pay a unit price $\bar{s}$ for the traffic.

With the sponsored data option, content providers choose which tariff applies, and this information is transmitted to consumers. Note that there is no possibility of discriminating between different content providers without inducing differential consumption. Indeed, if consumers were not affected by content providers' choices, then all such content providers would invariably opt for the same option. However, the network may attempt to increase its profits and the value it offers to consumers by combining a higher price to content providers with a zero price to consumers. Content providers eager to generate high traffic (stemming from high benefits) may be willing to choose this option, which corresponds to the sponsored data option.

In the case of sponsored data, if the network succeeds at inducing the LB and the HB content providers to choose different tariff options, then consumer behavior will adapt to the tariff, and the consumers will consume $\bar{q}$ of the sponsored content. The choice of the data allowance $T = (1-\lambda)\,\theta q_\ell$ amounts to choosing the consumption q_ℓ of LB content. By the same reasoning as in the previous section, the consumption of LB content is $q_\ell = D\,(r\theta)$, and the data allowance is $T = (1-\lambda)\,\theta D\,(r\theta)\,.$

We thus define the pricing under the *sponsored data option* as a menu $\{r, s, \bar{s}\}$ and consumption levels $\{q_\ell = D(r\theta), q_h = \bar{q}\}$ such that

- the LB and HB content providers are willing to participate, and
- the LB content providers choose the base tariff, while the HB content providers choose to sponsor the traffic.

Thus, the implicit price of the data allowance r determines the traffic of the LB content, while price s and $\bar{s}$ are used to screen the content. The consumer surplus is $\text{CS} = \lambda u(\bar{q}) + (1-\lambda) u(q_\ell)$, and the network profit is $\Pi_t = \lambda(\bar{s}-1)\bar{q}\theta + (1-\lambda)(s-1)\theta q_\ell$. The network maximizes the average joint surplus with consumers as follows:

$$V = \lambda\left[u(\bar{q}) + (\bar{s}-1)\theta\bar{q}\right] + (1-\lambda)\left[u(q_\ell) + (s-1)\theta q_\ell\right].$$

The resulting tariff induces participation by both HB and LB provider types as long as

$$b_\ell \geq s \text{ and } b_h \geq \bar{s}. \tag{3}$$

The following incentive compatibility conditions ensure that each content provider chooses the tariff designed for its specific type:

$$\begin{aligned} (b_\ell - s)\,\theta q_\ell &\geq (b_\ell - \bar{s})\,\theta\bar{q} \\ (b_h - \bar{s})\,\theta\bar{q} &\geq (b_h - s)\,\theta q_\ell \end{aligned} \tag{4}$$

Sponsored tariffs are thus equivalently characterized by an allocation $(q_l, \bar{q}, s, \bar{s})$ such that conditions (3) and (4) are both satisfied. The network's program is then to maximize V under these two constraints. This program departs from classical textbook cases because here the transfer $s\theta q$ depends on the quantity. Nevertheless, one can follow the usual procedure for solving such programs, which we describe next.

First, since it is optimal to raise content prices as long as they remain compatible with the constraints, if follows that the participation constraint of the LB content providers and the incentive compatibility constraint of the HB content providers will be binding, that is

$$s = b_\ell \quad \text{and} \quad (b_h - \bar{s})\,\bar{q} = (b_h - s)q_\ell. \tag{5}$$

Second, under condition (5), all constraints are satisfied because $q_\ell \leq \bar{q}$. If we replace the prices with the values given by condition (5), the reduced program can be written as a function of the quantities:

$$\max_{q_\ell} \lambda\left[u(\bar{q}) - (1-b_h)\theta\bar{q} - (b_h - b_\ell)\theta q_\ell\right] + (1-\lambda)\left[u(q_\ell) - (1-b_\ell)\theta q_\ell\right].$$

This expression leads directly to the following statement.

Proposition 3 *When sponsored data is offered, the network tariffs are*

$$r^* = \min\left\{1 - b_\ell + \frac{\lambda}{1-\lambda}(b_h - b_\ell),\ \frac{\bar{p}}{\theta}\right\};\ s^* = b_\ell,\ \bar{s}^* = b_h - (b_h - b_\ell)\,\frac{q_\ell^*}{\bar{q}},$$

with consumption levels $q_\ell^* = D\left(r^*\theta\right)$ *and* $q_h^* = \bar{q}$.

Proof. The solution of the reduced program is obtained at

$$\begin{aligned} u'\left(q_\ell\right) &= r\theta = \left(1 - b_\ell\right)\theta + \frac{\lambda}{1-\lambda}(b_h - b_\ell)\theta \text{ if } r \text{ is less than } \bar{p} \\ q_\ell &= 0 \text{ otherwise} \end{aligned}$$

This determines the usage price for consumers. In the event that $q_\ell = 0$, we adopt the convention that $r^*\theta = \bar{p}$ for clarity, but any larger price would also work. The content prices are then given by condition (5). ■

The menu of tariffs proposed by the network plays two roles. It allows the network to screen the different types of content providers, and it leads to more efficient consumption than under uniform pricing. Consider the tariff designed for the HB content providers. Because the gains generated by these providers' traffic are higher than its cost, they prefer high consumption and choose the option that removes it from the consumer data allowance. This amounts to choosing a price equal to zero for consumers, thereby inducing an efficient consumption level of the HB content. The price $\bar{s}$ paid by the HB content providers is strictly less than b_h because the network must leave some profit to induce the HB content providers to choose the sponsored tariff option. Whereas the price s paid by the LB content providers is simply set to minimize their profit, the price r paid by consumers to access that content is affected by two factors. First, r reflects the net cost of any unit of consumption. Second, r is distorted so as to minimize the profit that the network must leave to the HB content providers and still induce them to choose the right tariff.[9]

In this setting, the network may decide to exclude LB content providers. This is the case when the quantity q_ℓ resulting from the price characterized in Proposition 3 is negative — and thus when λ is large:

$$q_\ell = 0 \iff \lambda \geq \lambda^* = \frac{\bar{p} - (1-b_\ell)\theta}{\bar{p} - (1-b_h)\theta}. \tag{6}$$

[9] This bears some similarities to the analysis of capacity choices by Choi and Kim (2010). In their paper, as in ours, the network reduces that value of the basic services to the website to raise the price of the premium service.

In this case, the network may simply rely on a uniform tariff $s = b_h$ and $r = 0$. Given a low proportion of HB content, the network chooses to induce full participation and screens between the two content types. Thus, the sponsored data option is preferred to uniform pricing for $\lambda < \lambda^*$.

Corollary 1 *When sponsored data is allowed, the network proposes this option when $\lambda < \lambda^*$ and excludes the LB content otherwise. There is less exclusion than without sponsored data ($\lambda^u < \lambda^*$). As in the case of uniform pricing, the threshold for exclusion (now λ^*) is decreasing in b_h and increasing in b_ℓ.*

Proof. See Appendix. ■

3.3 Ban on sponsored data

Let us now consider the welfare consequences of banning sponsored data. This could come from a regulatory rule such as that in the net neutrality debate. As there is always exclusion of the LB content when $\lambda > \lambda^*$, we focus on the case in which the fraction of HB content is not too large. We assume that the regulator seeks to maximize total welfare. In the case of sponsored data, this welfare is

$$W^* = \lambda \left[u\left(\bar{q}\right) + \left(b_h - 1\right) \theta \bar{q} \right] + \left(1 - \lambda\right) \left[u\left(q_\ell^*\right) + \left(b_\ell - 1\right) \theta q_\ell^* \right].$$

Since under sponsored data, q_h^* is efficient while $q_\ell^* < q_\ell^{\mathrm{FB}}$, the main regulatory concern will be to increase consumption of LB content while avoiding any decrease in the level of HB content consumption.

We have shown in Section 3 that there will be more exclusion under uniform pricing because the network cannot accommodate the LB content, which should generate only low levels of consumption, and simultaneously exploit its market power on the HB content. However, a network that wants to attract users who consume both types of content must raise the level of LB consumption and reduce the level of HB consumption. Thus, as is often the case with price discrimination, the overall effect of a ban on sponsored data is ambiguous.

When there is no exclusion, social welfare under uniform pricing is given by

$$W^u = u\left(q_\ell^{\mathrm{FB}}\right) + \lambda \left(b_h - 1\right) \theta q_\ell^{\mathrm{FB}} + \left(1 - \lambda\right) \left(b_\ell - 1\right) \theta q_\ell^{\mathrm{FB}}.$$

Sponsored data yields more efficient consumption of the HB content but less efficient consumption of the LB content. As the proportion λ of HB content increases, the gain on HB content becomes more valuable, but the distortion on the LB content increases. Therefore, the overall comparison is ambiguous. We can nevertheless state the following result.

Proposition 4 *If sponsored data is banned, then*

- *when $\lambda^* > \lambda > \lambda^u$, total welfare decreases;*
- *when $\lambda < \lambda^u$, total welfare decreases if λ is small enough and increases when b_ℓ is close to 1 (i.e., a_l close to the cost θ).*

Proof. See Appendix ■

When $\lambda^* > \lambda > \lambda^u$, allowing sponsored data benefits both groups of content providers. The LB content providers are now active, whereas they were excluded before. The HB content providers also benefit from the introduction of the sponsored data option because, to preserve incentive compatibility, the network must leave them some rent — which is not the case when uniform prices are exclusionary.

When there is no exclusion under either regime (for $\lambda < \lambda^u$), the effect of a ban on sponsored data is more ambiguous. In that case, there is little consumption of HB content whereas the consumption of LB content increases to the efficient level, and welfare could either rise or fall as a result. When λ is small, the distortion of q_ℓ (or, equivalently, of r) under sponsored data is likewise small, and thus, the former effect dominates. However, when b_l is close to 1, the distortion of q_ℓ under uniform pricing small, such that the latter effect dominates.

In short, sponsored data induces some changes in the price and therefore the consumption of both types of content. When the price under uniform pricing is high, or the quantity consumed is low, sponsored data reduces significantly the price of HB content, and this is socially profitable. When the price under uniform pricing is low, or the quantity consumed is high, sponsored data's major impact is to increases the price of LB content to allow the network to decrease the rent of the HB content providers, and social welfare is decreased.

4 Extensions

4.1 One-sided pricing

A possible interpretation of net neutrality regulation is that it consists of imposing one-sided pricing, which is a *zero-price rule* for content (see, for example, Economides and Täg (2012)). As the current situation is similar to one-sided pricing, it is worth investigating the implications of sponsored data in this setting. Note that because sponsored data relates to consumers' data allowance and not directly to traffic, there is no inconsistency in allowing it in the context of a zero price being charged to content providers for traffic. Therefore, the regulator may or may not allow it.

Thus, let us assume that the price s is constrained to be equal to zero. When sponsored data is not possible, the network sets a unique consumer price $r = 1$ that maximizes the network value $u(D(r\theta)) - \theta D(r\theta)$, which is then

$$V_0^u = u(D(\theta)) - \theta D(\theta).$$

and the consumption of any content is given by $q_h = q_\ell = D(\theta)$. In this situation, the network cannot exclude any content and thus chooses a high usage price for consumers.

Let us suppose now that sponsored data is allowed. This means that the network chooses a tariff $P(T)$ along with a zero price ($s = 0$) for content providers and may offer content providers the option of sponsoring consumption at price $\bar{s}$ per unit of content. In this case, the network's pricing program is the same as in Section 3 except that the constraint $s \leq b_\ell$ is replaced by the new constraint $s = 0$. The reasoning of Proposition 3 applies with this new constraint, which leads to the following optimal strategy for the network:

$$r = \min\left(1 + \frac{\lambda}{1-\lambda} b_h, \frac{\bar{p}}{\theta}\right) > r^*; \qquad \bar{s} = b_h - b_h \frac{q_\ell}{\bar{q}}. \tag{7}$$

To induce the HB content providers to choose the sponsoring option with a positive price $\bar{s}$ rather than the standard contract with a zero price, incentive compatibility of the offer requires the network to distort consumption of the LB content even more than the case of two-sided pricing ($q_\ell = D(r\theta) < q_\ell^*$). The network value under sponsored data is then

$$V_0^* = \lambda\left(u(\bar{q}) + (b_h - 1)\theta\bar{q} - b_h\theta q_\ell\right) + (1-\lambda)\left(u(q_\ell) - \theta q_\ell\right)$$

Because the network can offer sponsoring while maintaining the standard price of the data allowance at $r = 1$, it is always optimal to do so. However, the welfare implications are ambiguous.

Proposition 5 *Under one-sided pricing, the network always proposes a sponsored data option if it is allowed. A ban on sponsored data reduces total welfare when λ is large. When λ is small, the ban reduces welfare if b_ℓ is small but increases welfare if b_ℓ is large and demand is elastic enough.*

Proof. See Appendix ■

When a sponsored option is proposed, the consumption of HB content is unchanged, but the price of the data allowance is higher, and thus, the consumption of LB content is reduced. The attractiveness of sponsored data for content providers is naturally reduced when the network is not allowed to charge content providers for traffic. Hence, the network must reduce the non-sponsored consumption still further to maintain the sponsored data option's value. From a welfare perspective, sponsored data is better than uniform pricing when λ is large because there is efficient consumption of the HB content. When λ is small, the situation is more complex since the LB content consumption is only marginally distorted with sponsored data. However, the cost of this distortion depends on the benefit of the LB content. When this benefit is low, the consumption distortion is not excessively value-destructive, and the benefit from efficient consumption of the HB content dominates. When this benefit is high and demand is elastic enough, the low revenue due to reduced level of consumption outweighs the benefit from efficient consumption of HB content. Thus, the results obtained under two-sided pricing extend to the case of one-sided pricing.

4.2 Elastic participation and competition between networks

In the main analysis, we considered the case of a monopoly network with inelastic subscription demand. We now show that introducing demand elasticity or competition at the network level does not change the manner in which the variable cost is allocated between consumers and content providers — and thus does not affect the main conclusions of our work. For this purpose, let us describe in greater detail the participation decision of the consumers.

We consider a model with an initial unit mass of consumers, a unit mass of content providers and $I \geq 1$ networks (indexed by i). Content providers are further divided into a mass λ of type h and a complementary mass $1-\lambda$ of type ℓ. The utility of each consumer subscribing to network i and choosing a consumption profile $\{q_{ih}, q_{i\ell}\}$ and data allowance T_i is given by[10]

$$\lambda u(q_{ih}) + (1-\lambda)\, u\,(q_{i\ell}) - F_i - r_i T_i + \tilde{\varepsilon}_i$$

where F_i is the hookup fee, r_i the price of the data allowance, and $\tilde{\varepsilon}_i$ is an idiosyncratic shock. The idiosyncratic shock $\tilde{\varepsilon}_i$ is a random variable that represents consumers' heterogeneity with respect to the intrinsic taste for network i. We impose no restrictions on the distribution of preference shocks, but we implicitly assume that they do not convey any information about the utility derived from consuming content.[11]

The timing of the game is unchanged, and we assume that in stage 1, each network i si-

[10] If type-t content is not available on the network, we set $q_{it} = 0$.

[11] This modeling of competition can be seen as a simplified version of the "nested discrete choice" model of demand developed in Anderson and de Palma (1992).

multaneously makes an offer $(F_i, r_i, s_i, \bar{s}_i)$. In this slightly modified setting, content providers may deliver their content to all networks — they then pay only a variable price — whereas consumers subscribe to a single network.

Let N_i denote the mass of consumers subscribing to network i. Let $s_{it} \in \{s_i, \bar{s}_i\}$ be the price paid by a content provider of type t. Then, the profit of each content provider using this network is given by

$$N_i(a_t - s_{it}\theta)q_{it} = N_i\theta(b_t - s_{it})q_{it}.$$

A content provider of type t will choose to participate in network i if $b_t \geq \min\{s_i, \bar{s}_i\}$. In this context, the participation of content providers in network i and their choice of tariffs, as well as individual-level consumption for a given contract, are the same as before. What differs is that consumers can now choose among networks.

The gross consumer surplus is given by

$$\text{CS}_i = \lambda u\left(q_{ih}\right) + (1-\lambda)\, u\left(q_{i\ell}\right).$$

A given consumer joining network i gains $\text{CS}_i + \varepsilon_i - F_i - r_i T_i$. Because there are several networks, the mass of consumers subscribing to network i is given by

$$N_i = \Pr\left(\text{CS}_i - F_i - r_i T_i + \varepsilon_i \geq \max\{0, \max_{j \neq i} \text{CS}_j - F_j - r_j T_j + \varepsilon_j\}\right).$$

The total profit of network i is then

$$N_i\left[F_i + r_i T_i + \lambda(s_{ih} - 1)\theta q_{ih} + (1-\lambda)\left(s_{i\ell} - 1\right)\theta q_{i\ell}\right].$$

For any given strategy of the other networks (denoted z_{-i}), let

$$\phi_i\left(R; z_{-i}\right) = \Pr\left(R \geq \max\{0, \max_{j \neq i} CS_j - F_j - r_j T_j + \varepsilon_j\} - \varepsilon_i\right).$$

Now, we can write the profit of network i as

$$\phi_i\left(\text{CS}_i - F_i - r_i T_i; z_{-i}\right)\left[F_i + r_i T_i + \lambda(s_{ih} - 1)\theta q_{ih} + (1-\lambda)\left(s_{i\ell} - 1\right)\theta q_{i\ell}\right].$$

Under this formulation, it is easy to see that the networks' best pricing strategy always maximizes the network value per consumer.

Proposition 6 *In any equilibrium of the game with elastic subscription demand and I networks, each network chooses a tariff structure $(r_i, s_i, \bar{s}_i)$ that maximizes its value per con-*

sumer: $V_i = \lambda \left(u\left(q_{ih}\right) + \left(s_{ih} - 1\right)\theta q_{ih}\right) + \left(1-\lambda\right)\left(u\left(q_{i\ell}\right) + \left(s_{i\ell} - 1\right)\theta q_{i\ell}\right)$. *Hence, the optimal tariff structure* $(r^*, s^*, \bar{s}^*)$ *is not affected by demand elasticity or competition, only the hook-up fee* F^* *is.*

Proof. Let $R_i = \mathrm{CS}_i - F_i - r_i T_i$ be the expected net consumer surplus and V_i be the network value defined above. The network's profit can be written as

$$\phi_i\left(R_i; z_{-i}\right)\left[V_i - R_i\right].$$

Note that V_i is independent of the subscription fee F_i and of the other networks' strategies z_{-i}. Indeed, V_i depends solely on $(r_i, s_i, \bar{s}_i)$ through their effect on quantities and revenue per user from content providers.[12] Let us fix the participation $\phi_i\left(R_i; z_{-i}\right)$ by adjusting F_i to maintain R_i constant. Then, profit maximization for a given R_i implies that the network will always choose $(r_i, s_i, \bar{s}_i)$ to maximize V_i. ■

The value V_i is solely dependent on the usage prices, meaning that is a natural ordering in the pricing strategy. First, the network maximizes the value that can be shared with consumers by setting adequate usage prices. Then, the network decides how much surplus to retain and how much surplus to leave to the consumers. Whereas the surplus R_i left to consumers (and hence the subscription fee F_i) depends on the elasticity of demand and on competition between networks, the prices $(r_{it}, s_{it})_{t\in\{\ell,h\}}$ do not. It follows that the prices derived in the main model with a monopoly network are also the equilibrium prices in the case of many networks competing for consumers.

As far as welfare is concerned, if total demand is fixed (inelastic consumer participation), then introducing competition at the network level does not alter our results. However, when aggregate demand is elastic, competition may increase total participation in the market. We observe that, compared with the case of inelastic demand, the regulation of the traffic prices should be more favorable to sponsored data under competition.

Corollary 2 *Sponsored data is all the more welfare-enhancing compared with uniform pricing as aggregate demand becomes more elastic.*

Proof. See Appendix. ■

4.3 Heterogenous cost and demand

We can introduce other dimensions of heterogeneity by allowing the content providers to differ in their cost to the network and in the demand for their product. More precisely, we

[12] More generally, V_i depends on the slope of the tariff P_i but not on its level.

now assume that content providers' types differ in terms of both θ and a. Content of type h is characterized by (θ_h, a_h), while content of type ℓ is characterized by (θ_ℓ, a_ℓ), with probability λ and $1-\lambda$, respectively. We still rank these two types of content providers according to their respective ratios of advertising revenue to cost and continue to assume that

$$b_\ell = \frac{a_\ell}{\theta_\ell} < 1 < b_h = \frac{a_h}{\theta_h}.$$ [13]

We also allow the demands to differ across content types by assuming that the utility derived from the consumption Q of type-t content is $\eta_t u\left(Q/\eta_t\right)$. Although this is not essential to the argument, we assume for simplicity that the utility function is quadratic and thus that the demand function D is linear.

For the sake of comparison with the previous section, we denote by Q_t the consumption of type t content and by $q_t = Q_t/\eta_t$ the "normalized" consumption. In this setting,

i) $Q_t = \eta_t D\left(r\theta_t\right)$ $\left(q_t = D\left(r\theta_t\right)\right)$ if the content is available, i.e., if s is below b_t;

ii) $Q_t = q_t = 0$ if it is not available., i.e., s is above b_t.

The analysis under uniform pricing is then the same as before, adjusting for the new demand. The network must choose between two prices for the content providers, $s = b_\ell$ and $s = b_h$.

In the case in which $s = b_\ell$, the price for data charged to consumers is $r = 1 - b_\ell$. The value captured by the network is

$$\begin{aligned} V_\ell^u &= \lambda\eta_h\left(u\left(q_h^u\right) + \left(b_\ell - 1\right)\theta_h q_h^u\right) + \left(1-\lambda\right)\eta_\ell\left(u\left(q_\ell^u\right) + \left(b_\ell - 1\right)\theta_\ell q_\ell^u\right) \\ \text{where } q_t^u &= Q_t^u/\eta_t = D\left(\left(1-b_\ell\right)\theta_t\right). \end{aligned}$$

In the case in which $s = b_h$, then the price for data is $r = 0$, the consumer consumes a quantity $\eta_h \bar{q}$ of HB content, and the LB content is excluded. The value is

$$V_h^u = \lambda\eta_h\left(u\left(\bar{q}\right) + \left(b_h - 1\right)\theta_h\bar{q}\right)$$

In this case again, there is a critical value λ^u such that exclusion occurs if $\lambda > \lambda^u$.

Consider now the introduction of a sponsored data option at unit price $\bar{s}$. Then a type-t content provider will opt for this option if $\left(b_t - \bar{s}\right)\theta_t\eta_t\bar{q} \geq \left(b_t - s\right)\theta_t\eta_t D\left(r\theta_t\right)$, which reduces to

$$\left(b_t - \bar{s}\right)\bar{q} \geq \left(b_t - s\right)D\left(r\theta_t\right)$$

[13] The model is also compatible with a uniform benefit $a_\ell = a_h$ per unit of consumption across types and different costs associated with consumption, in which case type h is low-cost content.

From the network's perspective, the two main benefits of sponsoring are that it enables the extraction of more rent from content providers and can induce more efficient levels of consumption. This interaction between screening on one side and incentives on the other side is a distinguishing feature of sponsored data in this extended setting. The optimal sponsored pricing mechanism is obtained as before but with a new objective, which is written as follows:

$$\lambda \left[\eta_h u\left(q_h\right) + (s_h - 1)\theta_h \eta_h q_h\right] + (1-\lambda)\left[\eta_\ell u\left(q_\ell\right) + (s_\ell - 1)\eta_\ell \theta_\ell q_\ell\right].$$

with s_ℓ and s_h in $\{s, \bar{s}\}$. The characterization of the solution can then be extended as described next.

Proposition 7 *Suppose that content types differ in terms of costs, benefits and demands. Then, there exists $\lambda^* < 1$ such that the following hold:*

a) *For $\lambda < \lambda^*$, the network's profit-maximizing sponsored tariffs are such that the HB content providers choose to sponsor consumption and*

$$\begin{aligned} s^* &= b_\ell,\ r^* = 1 - b_\ell + \frac{\lambda}{1-\lambda}(b_h - b_\ell)\frac{\eta_h}{\eta_\ell}\left(\frac{\theta_h}{\theta_\ell}\right)^2 < \frac{\bar{p}}{\theta_\ell} \\ \bar{s}^* &= b_h - (b_h - b_\ell)\frac{D\left(r^* \theta_h\right)}{\bar{q}}. \end{aligned}$$

b) *When $\lambda > \lambda^*$, the HB content providers pay $\bar{s}^* = b_h$ for consumption $\bar{q}$ and*

- **i)** *If $\theta_\ell \geq \theta_h$, there is exclusion of the LB content*;
- **ii)** *If $\theta_h > \theta_\ell > \left(1 - \frac{q_\ell^{\mathrm{FB}}}{\bar{q}}\right)\theta_h$, $q_\ell = \bar{q}\left(1 - \frac{\theta_\ell}{\theta_h}\right) > 0$.*
- **iii)** *If $\left(1 - \frac{q_\ell^{\mathrm{FB}}}{\bar{q}}\right)\theta_h > \theta_\ell$, $q_\ell = q_\ell^{\mathrm{FB}}$ for all $\lambda \geq 0$ (and $\lambda^* = 0$).*

c) *Except when exclusion occurs, sponsored data is more profitable than uniform pricing.*

Proof. See Appendix. ■

The new feature is that there is no exclusion if under equal consumption, the LB content induces lower traffic than the HB content. In this case, one can find a price of the data allowance that deters consumption of HB content but not of LB content. Thus, the network needs not to foreclose the LB content to gain maximal profit from the HB content providers.

As before, exclusion may be achieved by way of a uniform tariff $s = b_h$. When the proportion of HB content is below λ^* or when $\theta_\ell < \theta_h$, the network accommodates the LB content via sponsoring HB content. The consumption of the LB content is then sub-optimal.

A lesson from the result is that the distinction between a small benefit b and a large benefit b is the relevant one to determine which content provider will choose the sponsored data option. What matters is thus the return on traffic (the revenue for each unit of traffic). In particular, the size of demand affects the prices but not which content providers will sponsor data.

Proposition 8 *Suppose that content types differ in terms of costs, benefits and demands. If sponsored data is banned, then*

- *when $\lambda^* > \lambda > \lambda^u$ or when $\lambda > \lambda^*$ but $\theta_h > \theta_\ell$, total welfare decreases;*
- *when $\lambda < \lambda^u$, total welfare decreases if λ is small enough and increases if b_h is close to 1.*

Proof. see Appendix. ■

The welfare comparison is similar to the case in which only benefits are heterogenous. A new effect is that sponsored data yields a further social benefit in that it avoids complete exclusion of the LB content when the proportion of HB content is large.

Similarly, a zero-price regulation reduces welfare if b_h is high because in that case there will be insufficient (resp. excessive) consumption of the HB (resp. LB) content.

4.4 Paid and free content

In the paper we have contrasted the case where there is only paid content with the case where there is only free content. In Appendix B, we analyze the situation in which paid content providers and free content providers coexist, and we provide here the main insights of this analysis.

We assume that there is a fraction β of paid content and a fraction $1-\beta$ of free content. The characteristics of the population of free content providers is as before. The paid content providers have no benefit ($a = 0$) and face a demand $\hat{D}$ derived from utility $\hat{u}$. From section 2.2, we know that the contribution of a paid content to the network value is

$$V_p(w) = \hat{u}(\hat{D}(\hat{p}^m(w\theta))) - \hat{p}^m(w\theta)\hat{D}(\hat{p}^m(w\theta)) + (w-1)\theta\hat{D}(\hat{p}^m(w\theta)),$$

where $w = r + s$ is the total price paid for traffic. This value is maximal at w_p^*.

The contribution of a type-t free content when active at prices r and s is

$$V_f(r,s) = u(D(r\theta)) + (s-1)\theta D(r\theta).$$

We assume that both V_p and V_f are concave.

The analysis of this case is rather complex because the impact of prices on free and paid content differs and may lead to conflicting effects. In this discussion we focus on the case where $w_p^* > b_h$, which means that the total price the network would like to set for the paid content if it could discriminate is larger than the price it would charge for the HB free content.[14]

When $w_p^* > b_h$, the price charged to the paid content providers is always too low from the network's perspective. As *(i)* the network is constrained on the content side of the market if it wants to keep some free content on board and *(ii)* it is always optimal for the network to accommodate the HB content,[15] the presence of paid content providers leads the network to inflate the consumer price of data allowance. This logic applies to uniform pricing and to sponsored data.

The value under uniform pricing is now

$$V = \beta V_p (s + r) + (1 - \beta) M (s) V_f (r, s) .$$

We can show that if λ is below a threshold $\hat{\lambda}^u$, the network charges $s = b_\ell$ and $r \geq 1 - b_\ell$, while if λ is above $\hat{\lambda}^u$, it charges $s = b_h$ and $r \geq 0$.[16]

The conclusions are thus similar to the case of free content only. The network chooses between extracting all rents from the LB content providers or excluding them and extracting all the rents from the HB content providers. The difference lies in the choice of the consumer price r. The network would benefit from charging a high total price for paid transactions, which may induce it to charge a higher price r to consumers than without the paid content.

The analysis of the value maximizing sponsored data is more involved as the solution depends on whether the paid content providers choose whether to sponsor. But still, the main conclusions are similar to the conclusions of the main model of the paper.

A first important result is that when $w_p^* > b_h$, it is never optimal to offer a sponsored data program and to exclude the LB content. As a consequence, when the network offers a sponsored data program, it always sets[17]

$$s = b_\ell \text{ and } \bar{s} = b_h - (b_h - b_\ell) \frac{D(\theta r)}{\bar{q}}.$$

The presence of paid content however affects the consumer price r and we show that again

[14] This is the case if the pass-through on paid content is not too large.

[15] The network can set $s = b_h$ and $r = w_p^* - b_h$ and do better than with exclusion of all free contents.

[16] If $w_p^* < b_h$, the network may choose to reduce the price s below b_h when excluding the LB content.

[17] If $w_p^* < b_h$, the network may choose to reduce the price s below b_ℓ.

this price is increased:

$$r > 1 + \frac{\lambda}{1-\lambda}\left(b_h - b_\ell\right).$$

Indeed, when the paid content providers choose not to sponsor, they pay $w = r + s$ and increasing r directly raises the value derived from them. When the paid content providers choose to sponsor instead, they pay a price $w = \bar{s}$ to the network. Then increasing r has no direct effect on w but, by relaxing the incentive constraint of the HB content providers, it allows the network to raise $\bar{s}$ and w. Thus, either directly or indirectly, raising the consumer price r allows to extract more rent from the paid content providers.

Overall, the presence of paid content leads to more eviction of LB content. Indeed exclusion becomes relatively more attractive as it allows raising the total price for paid content. More generally the consumption of LB content declines. Moreover, as in the case with free content only, a ban of sponsored pricing raises exclusion.

If exclusion doesn't occur, a ban on sponsored data reduces the consumer price r if β is small or if $\bar{q} + \left(b_h - b_\ell\right)\theta D'\left(r\theta\right) \leq 0$. It would thus raise the welfare for the LB content providers. An additional effect is that, when the paid content is not sponsored, the lower consumer price would also benefit the paid content providers.

However we show that a ban on sponsored data raises the consumer price r if the proportion λ of HB content is small and $\bar{q} + \left(b_h - b_\ell\right)\theta D'\left(\bar{r}\theta\right) > 0$ (still if exclusion doesn't occur). Under these conditions, the paid content providers choose to sponsor ($w = \bar{s}$) and $\partial w/\partial r = \partial \bar{s}/\partial r < 1$. By contrast under uniform pricing with no exclusion, we have $w = b_\ell + r$ and so $\partial w/\partial r = 1$. Thus it is less costly (in terms of lost value on LB content) for the network to raise w under uniform pricing than under sponsored data. There are two conflicting effects of the ban on the incentives to raise the price r. First as in the case of free content only, the ban eliminates the incentives to reduce the rent of the HB content providers by raising r, as they pay $s = b_\ell$. But the ban raises the marginal benefit on paid content of increasing r. When λ is small the latter effect dominates and therefore the ban raises the price r.

5 Conclusion

It is long established that consumption efficiency requires that consumers receive adequate signals about the costs and benefits they impose on trading partners. In a standard setting with paid goods, this is achieved by the market price. When the goods are free, as is often the case on Internet, an alternative way to induce efficient consumption behavior must be found. Network prices may substitute for the missing content price in this case. This line of reasoning suggests that there are social benefits from giving contracting freedom to ISPs.

However, this comes at a cost when it leads to the exploitation of upstream market power. In this context, our analysis of sponsored data has highlighted some interesting elements.

The first insight relates to the two-sided nature of the Internet. We show that by boosting consumption, sponsored data plans allow the screening of traffic-sensitive content. Sponsoring then benefits not only the sponsor but also the consumers on the other side of the market. As it exploits the interactions between the two sides, this form of discrimination differs from one-sided discriminatory practices.

Second, allowing networks to propose sponsored data plans mitigates inefficient under-consumption for some content (high-value content or content that would be excluded otherwise) but worsens it for other content (low-value content). In this respect, although the mechanism differs, the conclusion is similar to that reached for price discrimination in one-sided markets.

Our analysis then suggests that whether a regulation that bans sponsored data is optimal depends on specific characteristics of the market. Nevertheless, as long as the sponsored data option concerns relatively few content providers and there are many low-value content providers, such a regulation is not welfare-enhancing.

Since networks provide unique access to consumers, it follows that even competitive networks will induce under-consumption of content, meaning that our analysis applies equally well to a monopoly ISP or to an oligopoly. The conclusions are also valid when ISPs are subject to a one-sided price regulation.

This paper did not address anti-trust issues that sponsored data may raise if some large content providers compete with smaller content providers. In this case, there is a risk that ISPs favor some content providers over others, by setting tariffs that are attractive only to large content providers. This may provide a motive for capping the price that may be charged for sponsoring data to reduce barriers to entry.

Another related issue that should be explored concerns vertical integration by ISPs in the content market. ISPs may be tempted to exempt consumers from data restrictions when they consume their own integrated content, thereby placing non-integrated content at a disadvantage.

To conclude, note that whether the market should be viewed as one-sided or two-sided depends on whether the content is paid or free (see Rochet and Tirole, 2006). An interesting extension of our work would be to discuss the choice of content providers to be free or not. Endogenizing this choice would provide interesting new insights into the optimal regulation of this industry.

References

[1] Anderson, S., and A., de Palma, 1992. "Multiproduct Firms: A Nested Logit Approach", *The Journal of Industrial Economics*, 40 (3), 261-276.

[2] Armstrong, M., 2006. "Competition in Two-Sided Market", *The RAND Journal of Economics*, 37 (3), 668-691.

[3] Bourreau, M., Kourandi, F. and T. Valletti, 2015, "Net Neutrality with Competing Platforms", *Journal of Industrial Economics*, 63, 30-73.

[4] Caillaud, B., and B., Jullien, 2003. "Chicken &Egg: competition among intermediation service providers", *The RAND Journal of Economics*, 34(2), 309-328.

[5] Choi, J.P., and B.-C., Kim, 2010. "Net Neutrality and Investment Incentives", *The RAND Journal of Economics*, 41(3), 446–471.

[6] Choi, J.P., Jeon D.-S. and B.-C., Kim, 2015. "Asymmetric Neutrality Regulation and Innovation at the Edges: Fixed vs. Mobile Networks", mimeo, Toulouse School of Economics.

[7] Economides, N. and B., Hermalin, 2012. "The Economics of Network Neutrality", *The RAND Journal of Economics,* 43(4), 602-629.

[8] Economides, N. and B. Hermalin, 2015. "The strategic use of download limits by a monopoly platform", *The RAND Journal of Economics*, 46(2), 297-327.

[9] Economides N. and J. Tåg, 2012. "Net Neutrality on the Internet : A two-sided Market Analysis", *Information Economics and Policy*, 24, 91–104.

[10] Greenstein S., Peitz M. and T. Valletti, 2016. "Net Neutrality: A Fast Lane to Understanding the Trade-offs", *The Journal of Economic Perspectives*, 30(2), 127-149.

[11] Hermalin, B. and M.L. Katz, 2004. "Sender or Receiver: who should pay to exchange an electronic message", *The RAND Journal of Economics*, 35(3), 423-448.

[12] Hermalin, B. and M.L. Katz, 2009. "Information and the hold-up problem", *The RAND Journal of Economics*, 40, 215-248.

[13] Jeon, D.S., Laffont, J.J., and J. Tirole, 2004."On the Receiver-Pays Principle", *The RAND Journal of Economics*, 35, 85-110.

[14] Krämer, J. and L. Wiewiorra, 2012. "Net neutrality and congestion sensitive content providers: Implications for content variety, broadband investment, and regulation", *Information Systems Research* 23(4), 1303-1321.

[15] Peitz, M. and F. Schuett, 2015. "Net Neutrality and the Inflation of Traffic", Tilburg Discussion Paper 205-06, Tilburg University.

[16] Reggiani, C. and T. Valletti, 2016. "Net neutrality and innovation at the core and at the edges", International Journal of Industrial Organization, 45, 16-27.

[17] Rochet J.C. and J. Tirole, 2003. "Platform Competition in Two-Sided Markets", *Journal of the European Economic Association*, 1, 990-1029.

[18] Rochet J.C. and J. Tirole, 2006. "Two-sided market: A progress Report", *The RAND Journal of Economics*, 37(3), 645-667.

[19] Strategy Analytics Reports, 2013. *Unlimited to Tiered Data Plans: User Preference and Management Needs.*

A Appendix

Proof of Proposition 2

While V_ℓ^u does not depend on λ, V_h^u is linear in λ. At $\lambda = 0$, we have $V_\ell^u > V_h^u = 0$ (because demand is positive at price $(1 - b_\ell)\,\theta$), and at $\lambda = 1$, we have $V_\ell^u < V_h^u$ (because $b_\ell < b_h$). This implies that $V_\ell^u < V_h^u$ for λ above a threshold λ^u and $0 < \lambda^u < 1$.

We define the surplus $S\,(p) = \max_{q \geq 0} u\,(q) - pq$. The threshold λ^u solves

$$\lambda^u = \frac{S\,((1 - b_\ell)\,\theta)}{S\,(0) + (b_h - 1)\,\theta D\,(0)}$$

which is decreasing in b_h and increasing in b_ℓ.

Proof of Corollary 1

The first and second claims are immediate from Proposition 3 and equation (6). Let us focus on the last claim, that is, $\lambda^* > \lambda^u$.

By the convexity of $S(.)$ and $S\,(\bar{p}) = 0$,

$$S\,((1 - b_\ell)\,\theta) < S\,(0)\,\frac{\bar{p} - (1 - b_\ell)\theta}{\bar{p}}$$

hence

$$\lambda^u < \frac{S(0)}{S(0) + (b_h - 1)\,\theta D(0)} \frac{\bar{p} - (1 - b_\ell)\theta}{\bar{p}}$$

Using $S(0) < \bar{p}D(0)$, we have (since the RHS increases with $S(0)$)

$$\lambda^u < \frac{\bar{p}D(0)}{\bar{p}D(0) + (b_h - 1)\,\theta D(0)} \frac{\bar{p} - (1 - b_\ell)\theta}{\bar{p}} = \frac{\bar{p} - (1 - b_\ell)\theta}{\bar{p} + (b_h - 1)\,\theta} = \lambda^*.$$

■

Proof of Proposition 4

If $\lambda^* > \lambda > \lambda^u$, it is immediate that sponsored data is better than uniform pricing, as the consumption level of HB content is the same, and there is no consumption of LB content under uniform pricing.

We focus now on the case in which $\lambda < \lambda^u$. Under sponsored data, expected social welfare is given by

$$\begin{aligned} W^* - W^u &= \lambda \left[u(\bar{q}) + (b_h - 1)\,\theta\bar{q} - u\left(q_\ell^{\mathrm{FB}}\right) - (b_h - 1)\,\theta q_\ell^{\mathrm{FB}} \right] \\ &\quad - (1 - \lambda) \left[u\left(q_\ell^{\mathrm{FB}}\right) + (b_\ell - 1)\,\theta q_\ell^{\mathrm{FB}} - u(q_\ell^*) - (b_\ell - 1)\,\theta q_\ell^* \right]. \end{aligned}$$

At $\lambda = 0$, we have $W^* = W^u$ and $q_\ell^* = q_\ell^{\mathrm{FB}}$. Using the first-order conditions above we have the following:

$$\frac{\partial(W^* - W^u)}{\partial\lambda} \,|_{\lambda=0} = u(\bar{q}) + (b_h - 1)\,\theta\bar{q} - [u(q_\ell^u) + (b_h - 1)\theta q_\ell^u] > 0.$$

Therefore, at least for small values of λ, sponsored pricing dominates uniform pricing.

Suppose finally that b_ℓ is close to 1; then, q_ℓ^{FB} is close to $\bar{q}$, while $q_\ell^* < q_\ell^{\mathrm{FB}}$. Hence, $W^* - W^u < 0$.

■.

Proof of Proposition 5

The first part follows from the fact that for all $q_\ell < \bar{q}$, a sponsored data option yields higher value:

$$\lambda [u(\bar{q}) + (b_h - 1)\,\theta\bar{q} - b_h\theta q_\ell] + (1 - \lambda) [u(q_\ell) - \theta q_\ell] > u(q_\ell) - \theta q_\ell. \tag{8}$$

As far as the welfare effects are concerned, the differential between the welfare under sponsored pricing and strict zero-price regulation is given by

$$\begin{aligned}\Delta &= \lambda\left[u(\bar{q})+(b_h-1)\theta\bar{q}-u\left(q^0\right)-(b_h-1)\theta q^0\right] \\ &\quad -(1-\lambda)\left[u\left(q^0\right)+(b_\ell-1)\theta q^0-u(q_\ell)-(b_\ell-1)\theta q_\ell\right],\end{aligned} \tag{9}$$

where $q_\ell^{\mathrm{FB}} > q^0 = D(\theta) > q_\ell$. This is positive at $\lambda = 1$, hence the second part of the proposition. The last part is more intricate since the welfare differential vanishes at $\lambda = 0$ because q_ℓ converges to q^0. Note first that

$$\frac{\partial q_\ell}{\partial \lambda} = D'(r\theta)\frac{b_h\theta}{(1-\lambda)^2}$$

The slope of the welfare differential is

$$\frac{\partial \Delta}{\partial \lambda} = \left[u(\bar{q})+(b_h-1)\theta\bar{q}-(b_h-b_\ell)\theta q^0\right]-\left[u(q_\ell)+(b_\ell-1)\theta q_\ell\right]+(\lambda b_h+(1-\lambda)b_\ell)\theta D'(r\theta)\frac{b_h\theta}{(1-\lambda)^2}$$

Moreover,

$$\frac{\partial \Delta}{\partial \lambda}\mid_{\lambda=0}= \left[u(\bar{q})+(b_h-1)\theta\bar{q}-u\left(q^0\right)-(b_h-1)\theta q^0\right]+b_\ell b_h\theta^2 D'(\theta) \tag{10}$$

with the first term being positive and the second negative. The second term vanishes when $b_\ell = 0$ (while q^0 and q_ℓ are not affected by b_ℓ), meaning that the slope is positive in this case. Therefore, for λ close to 0, sponsored pricing dominates zero-price when the benefit of LB firms is small.

Note also that by concavity,

$$\left[u(\bar{q})+(b_h-1)\theta\bar{q}-u\left(q^0\right)-(b_h-1)\theta q^0\right] < \theta b_h(\bar{q}-q_0).$$

Hence, the slope $\frac{\partial\Delta}{\partial\lambda}\mid_{\lambda=0}$ is negative if

$$\frac{\bar{q}-q_0}{q_0} < b_\ell\left(-\frac{\theta D'(\theta)}{D(\theta)}\right),$$

therefore when demand is elastic and the LB content benefit is large.

Proof of Corollary 2

Consider first a monopoly network with elastic demand $\phi(\mathrm{CS}-F-rT)$. Let V^* and W^* be the network value and welfare per consumer with sponsored data. Let $\bar{V}$ and $\bar{W}$ be the network value with uniform pricing.

With network value V, the network chooses the fee F by solving

$$\begin{aligned} & \max_{F} \phi(R)(V-R) \\ R &= \mathrm{CS} - F - rT \end{aligned}$$

The first-order condition for F then yields

$$R + \frac{\phi(R)}{\phi'(R)} = V$$

which defines R as an increasing function of V. Thus, with elastic participation, the consumers' participation $\phi(R)$ is a function $N(V)$, increasing with V.

Hence participation N^* with sponsored data is higher than participation $\bar{N}$ under a uniform price. Whenever $N^*/\bar{N} > \bar{W}/W^*$, allowing sponsored data dominates. Whether this occurs depends on the elasticity of demand. Note that V^*, W^*, $\bar{V}$ and $\bar{W}$ are independent of ϕ; thus, the ratio $N^*/\bar{N}$ increases when demand becomes more elastic. Thus, allowing sponsored data will dominate for very elastic demand. ■

Proof of Proposition 7

We proceed in three steps. First, we derive the optimal contract when the h-type provider chooses the sponsored option. Then, we show that it is profitable to use this contract rather than the uniform contract. Finally, we show that it is not possible to make the l-type provider (and only this type) choose the sponsored option.

Let us first derive the optimal contract when the h-type provider chooses the sponsored option. The network's objective is

$$\max \lambda \left[\eta_h u(\bar{q}) + (\bar{s}-1)\theta_h \eta_h \bar{q}\right] + (1-\lambda)\left[\eta_\ell u(q_\ell) + (b_\ell - 1)\eta_\ell \theta_\ell q_\ell\right]$$

and the constraints

$$\begin{aligned} (b_h - \bar{s})\,\bar{q} &\geq (b_h - s)\, D(r\theta_h) \\ (b_\ell - s)\, D(r\theta_\ell) &\geq (b_\ell - \bar{s})\,\bar{q} \end{aligned}$$

along with

$$b_\ell \geq s$$

As we can increase $\bar{s}$, we have $(b_h - \bar{s})\,\bar{q} = (b_h - s)\,D\,(r\theta_h)$. The constraints are then

$$\begin{aligned} \bar{s}\bar{q} &= b_h\bar{q} + (s - b_h)\,D\,(r\theta_h) \\ b_h\bar{q} - b_h D\,(r\theta_h) + b_\ell D\,(r\theta_\ell) - b_\ell\bar{q} &\geq s\,(D\,(r\theta_\ell) - D\,(r\theta_h)) \end{aligned}$$

Suppose that $D\,(r\theta_\ell) \leq D\,(r\theta_h)$; then, we can increase s such that $s = b_\ell$. Suppose instead that $D\,(r\theta_\ell) > D\,(r\theta_h)$; then, we have

$$b_h\bar{q} - b_h D\,(r\theta_h) + b_\ell D\,(r\theta_\ell) - b_\ell\bar{q} \geq b_\ell\,(D\,(r\theta_\ell) - D\,(r\theta_h))$$

because $D\,(r\theta_h) \leq \bar{q}$. Again, we find that it is optimal to set $s = b_\ell$. Note that this also implies that the incentive constraint of the LB content provider is slack. Therefore, we have

$$s = b_\ell; \quad \bar{s}\bar{q} = b_h\bar{q} + (b_\ell - b_h)\,D\,(r\theta_h)$$

The optimum solves the following program (called (SP)):

$$\begin{aligned} &\max \lambda\left[\eta_h u\,(\bar{q}) + (b_h - 1)\,\theta_h\eta_h\bar{q} + (b_\ell - b_h)\,\theta_h\eta_h\tilde{q}_h\right] + (1-\lambda)\left[\eta_\ell u\,(q_\ell) + (b_\ell - 1)\eta_\ell\theta_\ell q_\ell\right] \\ &s.t. \quad q_\ell = D\,(r\theta_\ell)\,, \quad \tilde{q}_h = D\,(r\theta_h) \end{aligned}$$

or, equivalently,

$$\begin{aligned} &\max V_h^u + (1-\lambda)\,\eta_\ell\left(u\,(q_\ell) - (1-b_\ell)\theta_\ell q_\ell - \frac{\lambda}{(1-\lambda)}\,(b_h - b_\ell)\,\theta_h\frac{\eta_h}{\eta_\ell}\tilde{q}_h\right) \\ &s.t. \quad q_\ell = D\,(r\theta_\ell) \quad \text{and} \quad \tilde{q}_h = D\,(r\theta_h) \end{aligned}$$

In the case of a linear demand function, as $D(0) = \bar{q}$ and $D(\bar{p}) = 0$, we obtain $D(p) = \bar{q}(1 - \frac{p}{\bar{p}})$. Therefore, $q_\ell = \bar{q}\left(1 - \frac{r\theta_\ell}{\bar{p}}\right)$ and $\tilde{q}_h = \bar{q}\left(1 - \frac{r\theta_h}{\bar{p}}\right)$ such that for $q_\ell > 0$,

$$\tilde{q}_h = \max\left(\frac{\theta_h}{\theta_\ell}q_\ell + \bar{q}\left(1 - \frac{\theta_h}{\theta_\ell}\right), 0\right) > 0$$

We solve (after normalization)

$$\begin{aligned} &\max u\,(q_\ell) - (1-b_\ell)\theta_\ell q_\ell - \frac{\lambda}{(1-\lambda)}\,(b_h - b_\ell)\,\theta_h\frac{\eta_h}{\eta_\ell}\tilde{q}_h \\ st \quad \tilde{q}_h &= \max\left(\frac{\theta_h}{\theta_\ell}q_\ell + \bar{q}\left(1 - \frac{\theta_h}{\theta_\ell}\right), 0\right) \quad \text{if } q_\ell > 0 \\ \tilde{q}_h &= 0 \text{ if } q_\ell = 0 \ (exclusion) \end{aligned}$$

From the envelope theorem, the value is non-increasing in λ. Thus, there exist $\hat{\lambda}$ such that the value is zero and exclusion occurs if $\lambda > \hat{\lambda}$.

If $\theta_h \leq \theta_\ell$, we have $\tilde{q}_h \geq \frac{\theta_h}{\theta_\ell} q_\ell > 0$, and thus,

$$u'(q_\ell) = \theta_\ell \left(1 - b_\ell + \frac{\lambda}{1-\lambda}(b_h - b_\ell)\frac{\eta_h}{\eta_\ell}\left(\frac{\theta_h}{\theta_\ell}\right)^2\right) < \bar{p}$$

for $\lambda < \hat{\lambda} = \lambda^* < 1$, where at λ^* we have

$$u(q_\ell) - (1-b_\ell)\theta_\ell q_\ell - \frac{\lambda^*}{1-\lambda^*}(b_h - b_\ell)\frac{\eta_h}{\eta_\ell}\frac{\theta_h^2}{\theta_\ell} q_\ell = \frac{\lambda^*}{1-\lambda^*}(b_h - b_\ell)\frac{\eta_h}{\eta_\ell}\bar{q}\left(\theta_h - \frac{\theta_h^2}{\theta_\ell}\right).$$

Exclusion occurs when $\lambda > \lambda^*$.

If $\theta_h > \theta_\ell$, we have $\tilde{q}_h = 0$ when $q_\ell < \bar{q}\left(1 - \frac{\theta_\ell}{\theta_h}\right)$.

When $u'\left(\bar{q}\left(1 - \frac{\theta_\ell}{\theta_h}\right)\right) > \theta_\ell(1 - b_\ell)$, we find that $\hat{\lambda} = 1$ and

$$\begin{aligned} u'(q_\ell) &= \theta_\ell \left(1 - b_\ell + \frac{\lambda}{1-\lambda}(b_h - b_\ell)\frac{\eta_h}{\eta_\ell}\left(\frac{\theta_h}{\theta_\ell}\right)^2\right) \text{ if } \lambda < \lambda^*, \\ q_\ell &= \bar{q}\left(1 - \frac{\theta_\ell}{\theta_h}\right) < q_\ell^{\mathrm{FB}} \text{ if } \lambda > \lambda^*, \end{aligned}$$

where λ^* is the threshold where $\tilde{q}_h = 0$.

If $u'\left(\bar{q}\left(1 - \frac{\theta_\ell}{\theta_h}\right)\right) < \theta_\ell(1 - b_\ell)$, we have that $\tilde{q}_h = 0$ when $q_\ell = q_\ell^{\mathrm{FB}}$. In this case, the network can implement the first-best (q_ℓ^{FB} for LB and $\bar{q}$ for HB) and capture the entire surplus with $\bar{s} = b_h$ and $s = b_\ell$ for any value of λ.

Now, we can compare the profitability of this contract relative to the uniform contract. When there is no exclusion, the network value derived above is such that

$$\begin{aligned} &V^* = \max_r \lambda\left[\eta_h u(\bar{q}) + (b_h - 1)\theta_h \eta_h \bar{q} + (b_\ell - b_h)\theta_h \eta_h \tilde{q}_h\right] + (1-\lambda)\left[\eta_\ell u(q_\ell) + (b_\ell - 1)\eta_\ell \theta_\ell q_\ell\right] \\ &\textit{where.} \quad q_\ell = D(r\theta_\ell), \quad \tilde{q}_h = D(r\theta_h) \end{aligned}$$

Then, as $\eta_h u(\bar{q}) + b_h \theta_h \eta_h \bar{q} - \theta_h \eta_h \bar{q} > \eta_h u(\tilde{q}_h) + b_h \theta_h \eta_h \tilde{q}_h - \theta_h \eta_h \tilde{q}_h$:

$$\begin{aligned} V^* &> \max_r \lambda\left[\eta_h u(q_h) + b_\ell \theta_h \eta_h q_h - \theta_h \eta_h q_h\right] + (1-\lambda)\left[\eta_\ell u(q_\ell) + (b_\ell - 1)\eta_\ell \theta_\ell q_\ell\right] \\ &= V_\ell^u. \end{aligned}$$

This shows that it is more profitable for the network to offer a sponsored data option than a simple uniform contract.

Finally, we show that inducing the LB content producers to sponsor is not optimal.

First, separating the ℓ type is not possible. Indeed, suppose that we induce the ℓ type to sponsor. In this case, we must have

$$\begin{aligned}(b_h - \bar{s})\,\bar{q} &\leq (b_h - s)\,D\,(r\theta_h)\\(b_\ell - s)\,D\,(r\theta_\ell) &\leq (b_\ell - \bar{s})\,\bar{q}\end{aligned}$$

along with $b_\ell \geq \bar{s}$ for participation. As we can increase $\bar{s}$, we have

$$\bar{s} = \min\left(b_\ell + (s - b_\ell)\,\frac{D\,(r\theta_\ell)}{\bar{q}}, b_\ell\right)$$

Suppose that $s \geq b_\ell$. Then, $b_\ell = \bar{s}$ and we must have

$$(b_h - b_\ell)\,\bar{q} \leq (b_h - s)\,D\,(r\theta_h) \leq (b_h - s)\,\bar{q}$$

which is only possible if $s = b_\ell = \bar{s}$ and $r = 0$, which would be a suboptimal uniform tariff with no sponsored pricing.

Suppose that $s < b_\ell$. Then, $\bar{s} = b_\ell + (s - b_\ell)\,\frac{D(r\theta_\ell)}{\bar{q}}$, and we must have

$$(b_h - b_\ell)\,\bar{q} + (b_\ell - s)\,D\,(r\theta_\ell) \leq (b_h - s)\,D\,(r\theta_h)$$

which is not possible if $\theta_\ell \leq \theta_h$. In the case in which $\theta_\ell > \theta_h$, both prices s and $\bar{s}$ are below b_ℓ , which is dominated by a uniform price at $s = b_\ell$.

The last possibility is that both types sponsor the traffic. In this case, it is clearly optimal to set r very large and $\bar{s} = b_\ell$.

$$\lambda\left[\eta_h u\,(\bar{q}) + (b_\ell - 1)\theta_h\eta_h\bar{q}\right] + (1 - \lambda)\left[\eta_\ell u\,(\bar{q}) + (b_\ell - 1)\eta_\ell\theta_\ell\bar{q}\right],$$

which yields a lower payoff than V^* as is also obtained in program SP when $r = 0$.

Thus, it cannot be optimal to induce the type ℓ to sponsor. ■

Proof of Proposition 8

The proof is the same as the Proof of Proposition 4, adjusting for the new dimensions of heterogeneity. The welfare differential is

$$\begin{aligned}W^* - W^u &= \lambda\eta_h\left[u(\bar{q}) + (b_h - 1)\theta_h\bar{q} - u\left(q_\ell^{\mathrm{FB}}\right) - (b_h - 1)\theta_h q_\ell^{\mathrm{FB}}\right] \\ &\quad - (1-\lambda)\eta_\ell\left[u\left(q_\ell^{\mathrm{FB}}\right) + (b_\ell - 1)\theta_\ell q_\ell^{\mathrm{FB}} - u(q_\ell^*) - (b_\ell - 1)\theta_\ell q_\ell^*\right].\end{aligned}$$

The difference arises when $\lambda > \lambda^*$ and $\theta_h > \theta_\ell$ because there is no exclusion of LB content with sponsored pricing while there is without. Therefore, it is optimal to allow sponsored data. ■

B Paid and free content

Assume that there is a fraction β of pay content and a fraction $1-\beta$ of free content. The population of free content providers is as before. To simplify we assume pay content providers have no benefit ($a = 0$) and that the demand is $\hat{D}$ is derived from utility $\hat{u}$ such that the pass-through is less than 1

$$\begin{aligned}0 &< \frac{d\hat{p}^m(c)}{dc} < 1 \\ \textit{where } \hat{p}^m(c) &= \arg\max_p (p-c)\hat{D}(p)\end{aligned}$$

From section 2.2, we know that the contribution of a paid content to network value is, defining $\hat{q}^m(c) = \hat{D}(\hat{p}^m(c))$

$$V_p(w) = \hat{u}(\hat{q}^m(w\theta)) - \hat{p}^m(w\theta)\hat{q}^m(w\theta) + (w-1)\theta\hat{q}^m(w\theta)$$

where $w = r + s$ is the total price paid for traffic. Then V_p is maximal at $w_p^* > 1$.[18]

The contribution of a type-t free content when active at prices r and s is

$$V_f(r,s) = u(D(r\theta)) + (s-1)\theta D(r\theta).$$

where V_f is increasing in s, and maximal for given s at $r = \max\{1-s, 0\}$.

We assume that both V_p and V_f are concave.

[18] For instance, if $\hat{u} = \hat{q}q - \frac{q^2}{2}$ we obtain $\hat{q}^m(w\theta) = \frac{\hat{q}-w\theta}{2}$, $\hat{p}^m(w\theta) = \frac{\hat{q}+w\theta}{2}$ and $w_p^* = \frac{2\theta+\hat{q}}{3\theta}$.

B.1 Uniform pricing

Let us first derive the value maximizing tariffs. In the case of uniform tariff we have

$$V = \beta V_p(s+r) + (1-\beta) M(s) V_f(r,s)$$

where $M(s) = 1$ if $s \leq b_\ell$ and $M(s) = \lambda$ if $b_\ell < s \leq b_h$. We then obtain:

Proposition 9 *Under uniform pricing there is a threshold $\hat{\lambda}^u$ such that if $\lambda < \hat{\lambda}^u$, the network charge $s = b_\ell$ and $r \in \left(1 - b_\ell, w_p^* - b_\ell\right)$. If $\lambda > \hat{\lambda}^u$ it charges $r \geq 0$ and $s = b_h$ if $\beta \leq \hat{\beta} = \min\left\{\frac{\lambda\theta\bar{q}}{\lambda\theta\bar{q} - V_p'(b_h)}, 1\right\}$ and $w_p^* < s < b_h$ if $\beta > \hat{\beta}$.*

Proof. Let $\hat{r}(s) = \arg\max_{r\geq 0} \beta V_p(r+s) + (1-\beta) M(s) V_f(r,s)$.

Consider first $s \neq b_\ell$ and less than $\min\left\{b_h, w_p^*\right\}$. Because

$$\frac{\partial V}{\partial r} = \beta V_p'(r+s) + (1-\beta) M(s)(r+s-1)\theta^2 D'(r\theta)$$

we have $\max\{1-s, 0\} \leq \hat{r}(s) \leq w_p^* - s$. By the envelop theorem we have

$$\frac{dV(r(s),s)}{ds} = \beta V_p'(\hat{r}(s)+s) + (1-\beta) M(s)\theta D(\hat{r}(s)\theta) > 0.$$

Suppose that $b_h \leq w_p^*$. Then V increases on $0 \leq s < b_\ell$ and on $b_\ell \leq b_h$with a possible discontinuity at b_l. Moreover $s > b_h$ is not profitable because it is dominated by $s = b_h$ and $r = w_p^* - b_h$ (which yields maximum profit on paid content with sales to HB content). Thus in this case, the network chooses between $(s,r) = (b_\ell, r(b_\ell))$ and $(s,r) = (b_h, \hat{r}(b_h))$. The price $\hat{r}(b_h)$ may be positive or null.[19]

Suppose $b_h > w_p^*$, then the network sets $s = b_\ell$ or $w_p^* < s \leq b_h$. For $s > w_p^*$, the consumer price is $\hat{r}(s) = 0$ and

$$V = \beta V_p(s) + (1-\beta)\lambda V_f(0,s)$$

The network then chooses

$$s_{exc} = \arg\max_{b_\ell < s \leq b_h} \beta V_p(s) + (1-\beta)\lambda V_f(0,s).$$

Notice that the chosen s decreases with β and is equal to b_h for

$$\beta \leq \hat{\beta} = \frac{\lambda\theta\bar{q}}{\lambda\theta\bar{q} - V_p'(b_h)} < 1.$$

[19] We have $r(b_h) = 0$ if $\beta V_p'(b_h) + (1-\beta)\lambda(b_h - 1)\theta^2 D'(0) \leq 0$ which holds if β is small enough or b_h close to w_p^*.

While the value at $s = b_\ell$ is independent of λ, the value obtained when the LB content is excluded increases with λ. We then compare

$$\beta V_p \left(b_\ell + \hat{r}\left(b_\ell\right)\right) + \left(1-\beta\right) V_f \left(\hat{r}\left(b_\ell\right), b_\ell\right)$$

with

$$\max_{b_\ell < s \leq b_h} \beta V_p \left(\hat{r}\left(s\right) + s\right) + \left(1-\beta\right) \lambda V_f \left(\hat{r}\left(s\right), s\right)$$

Clearly the former dominate at $\lambda = 0$ because

$$\beta V_p \left(b_\ell + \hat{r}\left(b_\ell\right)\right) + \left(1-\beta\right) V_f \left(\hat{r}\left(b_\ell\right), b_\ell\right) \geq \beta V_p \left(w_p^*\right) + \left(1-\beta\right) V_f \left(w_p^* - b_\ell, b_\ell\right) > \beta V_p \left(w_p^*\right)$$

while the latter dominates when $\lambda = 1$. Thus there is a threshold below which the network chooses $s = b_\ell$. $\blacksquare$

When $b_h \leq w_p^*$ we have $\hat{\beta} = 1$. The conclusions are similar to the case of free content. Then the network chooses between extracting all rents from the LB content providers or excluding them and extracting all the rents from the HB content providers. The difference is in the choice of the consumer price r. When $b_h < w_p^*$, the network would benefit from charging a high total price for paid transactions (larger than the total price for any free content). This may lead the network to charge higher price r to consumers than without the paid content.

When $w_p^* \leq b_h$, the same logic applies for the non-exclusionary price $s = b_\ell$. However exclusion leads to a price larger than w_p^* so that the network will minimize the total price by setting $r = 0$ and may refrain from increasing the price up to b_h. The network may then leave a rent to the free HB content providers to avoid an excessive reduction of the volume of transactions of paid content.

B.2 Sponsored data.

Let us now turn to sponsored data, analyzing first the case in which there is no exclusion and then the case with exclusion.

B.2.1 No exclusion of the LB content providers

We suppose first that $s \leq b_\ell$ so that there is no exclusion. The program is given by

$$\max \beta V_p \left(w\right) + \left(1-\beta\right)\left(1-\lambda\right) V_f \left(r, s\right) + \left(1-\beta\right) \lambda V_f \left(0, \bar{s}\right)$$

s.t.

$$\begin{aligned} s &\leq b_\ell \ (IR_\ell); \ \bar{s} \leq b_h \ (IR_h) \\ w &= \min \{r+s, \bar{s}\} \quad (IC_p) \\ (b_h - \bar{s})\,\theta\bar{q} &\geq (b_h - s)\,\theta D(r\theta) \quad (IC_h) \\ (b_\ell - s)\,\theta D(r\theta) &\geq (b_\ell - \bar{s})\,\theta\bar{q} \quad (IC_\ell) \end{aligned}$$

Lemma 2 *When sponsored data is proposed, IC_h is binding and IC_ℓ is slack.*

If IC_h is slack, we must have $\bar{s} \leq s + r$ (otherwise the network would increase $\bar{s}$ until IC_h binds). In this case, the optimal prices would be $s = b_\ell$ and $r = 1 - b_\ell$ implying $\bar{s} < 1$ which cannot be optimal (again the network would raise $\bar{s}$). If $\bar{s} = s + r$, then the network can increases $\bar{s}$ without affecting w, s and r (which relaxes IC_ℓ). Thus IC_h binds which implies that IC_ℓ is slack. Using $\bar{s} = b_h - (b_h - s)\frac{q_\ell}{\bar{q}}$, we maximize

$$\begin{aligned} \tilde{V} &= \frac{\beta}{(1-\beta)(1-\lambda)} V_p(w) + u(D(r\theta)) + \left(s - 1 - \frac{\lambda}{1-\lambda}(b_h - s)\right)\theta D(r\theta) \\ &\quad + \frac{\lambda}{1-\lambda}\left(u(\bar{q}) + (b_h - 1)\,\theta\bar{q}\right) \end{aligned}$$

s.t.

$$s \leq b_\ell \ (IR_\ell);$$

$$w = \min\left\{r + s, b_h - (b_h - s)\frac{D(r\theta)}{\bar{q}}\right\} \quad (IC_p)$$

We then distinguish three types of solutions depending on the ranking between $r + s$ and $\bar{s}$.

1) Solution with $w = s + r < \bar{s}$ The slope of $\tilde{V}$ with respect to r is

$$\frac{\partial \tilde{V}}{\partial r} = \frac{\beta V_p'(s+r)}{(1-\beta)(1-\lambda)} + \left(s + r - 1 - \frac{\lambda}{1-\lambda}(b_h - s)\right)\theta^2 D'(r\theta)$$

which implies that the optimal $\hat{r}(s) = w(s) - s$ is such that

$$\min\left\{w_p^*, 1 + \frac{\lambda}{1-\lambda}(b_h - s)\right\} < s + r(s) < \max\left\{w_p^*, 1 + \frac{\lambda}{1-\lambda}(b_h - s)\right\}.$$

The slope of $\tilde{V}$ with respect to s is

$$\frac{\partial \tilde{V}}{\partial s} = \frac{\beta V_p'(s+r)}{(1-\beta)(1-\lambda)} + \frac{\theta D(\theta r)}{1-\lambda}$$

There are two subcases to analyze.

1.a) Case $s + r \leq w_p^*$

If we assume $w_p^* \geq b_h$ (the monopoly price is larger than the HB content willingness to pay), the condition $\bar{s} > s+r$ ensures that $s+r < w_p^*$ and thus that $\beta V_p'(w) + (1-\beta)\theta q_\ell > 0$ in the relevant range. Moreover $s + r < \bar{s}$ implies that $r < \bar{p}$. In this case an equilibrium where the paid content providers do not sponsor implies that

$$\begin{aligned} s &= b_\ell \\ r &= 1 - b_\ell + \frac{\lambda}{1-\lambda}(b_h - b_\ell) + \underbrace{\frac{-\beta V_p'(b_\ell + r)}{(1-\beta)(1-\lambda)\theta^2 D'(r\theta)}}_{>0} \\ \bar{s} &= b_h - (b_h - b_\ell)\frac{D(\theta r)}{\bar{q}} > b_\ell + r \end{aligned} \tag{11}$$

where the inequality ensures that $s + r < \bar{s}$.

More generally, given that the program is concave, the network sets $s = b_\ell$ if and only if $\beta V_p'(w(b_\ell)) + (1-\beta)\theta D(\theta r(b_\ell)) \geq 0$ which holds at least for

$$w_p^* \geq 1 + \frac{\lambda}{1-\lambda}(b_h - b_\ell).$$

1.b) Case $s + r > w_p^*$

If $w_p^* < 1 + \frac{\lambda}{1-\lambda}(b_h - b_\ell)$, then $V_p'(w(b_\ell)) < 0$ and there may exist some λ where $\beta V_p'(w(b_\ell)) + (1-\beta)\theta D(\theta r(b_\ell)) < 0$ at $s = b_\ell$ (recall that q_ℓ decreases in λ). In this case for λ large it may be optimal to reduce s below b_ℓ and to increases r. We thus obtain

$$\begin{aligned} s &\leq b_\ell, \\ r &= 1 - s + \frac{\lambda}{1-\lambda}(b_h - s) + \frac{-\beta V_p'(w)}{(1-\beta)(1-\lambda)\theta^2 D'(r\theta)}, \\ \bar{s} &= b_h - (b_h - s)\frac{D(\theta r)}{\bar{q}} > b_\ell + r. \end{aligned}$$

Note however that it must be the case that $b_\ell + r < b_h$ so that

$$\beta V_p'(w(b_\ell)) + (1-\beta)\theta D(\theta r(b_\ell)) > \beta V_p'(b_h) + (1-\beta)\theta D(\theta(b_h - b_\ell)).$$

It follows that this case cannot occur if

$$\beta < \tilde{\beta} = \min\left\{\frac{\theta D\left(\theta\left(b_h - b_\ell\right)\right)}{\theta D\left(\theta\left(b_h - b_\ell\right)\right) - V_p'\left(b_h\right)}, 1\right\}.$$

2) Solution with $s + r > \bar{s}$ It is immediate in this case that $s = b_\ell$ because only the LB content producers are affected by s. Then we have

$$\frac{\partial \tilde{V}}{\partial r} = -\left(\frac{\beta V_p'\left(\bar{s}\right)}{\left(1-\beta\right)\left(1-\lambda\right)} + \frac{\lambda}{1-\lambda}\theta\bar{q}\right)\left(b_h - b_\ell\right)\frac{\theta D'\left(r\theta\right)}{\bar{q}} + r\theta^2 D'\left(r\theta\right) + \left(b_\ell - 1\right)\theta^2 D'\left(r\theta\right)$$

which yields for an interior solution:

$$\begin{aligned} s &= b_\ell \\ r &= 1 - b_\ell + \frac{\lambda}{1-\lambda}\left(b_h - b_\ell\right) + \frac{\beta V_p'\left(\bar{s}\right)}{\left(1-\beta\right)\left(1-\lambda\right)\theta\bar{q}}\left(b_h - b_\ell\right) \\ \bar{s} &= b_h - \left(b_h - b_\ell\right)\frac{D\left(\theta r\right)}{\bar{q}} < b_\ell + r. \end{aligned} \tag{12}$$

3) Solution with $s + r = \bar{s}$. In this case the network could increase s and $\bar{s}$ for constant r and relax the incentives constraints. Given that $\bar{s} = b_h$ is not possible, we have $s = b_\ell$ along with

$$\begin{aligned} \bar{s} &= b_h - \left(b_h - b_\ell\right)\frac{D\left(r\theta\right)}{\bar{q}} \\ \bar{s} &= b_\ell + r \end{aligned}$$

or

$$r - \left(b_h - b_\ell\right)\left(1 - \frac{D\left(r\theta\right)}{\bar{q}}\right) = 0. \tag{13}$$

As $\left(b_h - b_\ell\right)\left(1 - D\left(0\right)/\bar{q}\right) = 0$, equation (13) admits $r = 0$ as a solution. But $r = 0$ cannot be optimal because then $\bar{s} = b_\ell$. Thus there cannot be any solution where $b_\ell + r = \bar{s}$ if equation 13 is monotonic in r, as it is the case for instance when D is a linear function.

Suppose that a solution $\bar{r} > 0$ to equation (13) exists (which requires that $r - \left(b_h - b_\ell\right)\left(1 - D\left(r\theta\right)/\bar{q}\right)$ is not monotonic) and that the value maximizing price is at $\bar{s} = s + \bar{r} < b_h$.

If $1 + \frac{\left(b_h - b_\ell\right)\theta D'\left(\bar{r}\theta\right)}{\bar{q}} > 0$ then increasing r with IC_h binding would yield $w = \bar{s} = b_h - \left(b_h - b_\ell\right)\frac{D\left(r\theta\right)}{\bar{q}} < b_\ell + r$, while decreasing r would yield $\bar{s} = b_h - \left(b_h - b_\ell\right)\frac{D\left(r\theta\right)}{\bar{q}} < w =$

$b_\ell + r$. The FOC conditions for r are then

$$\begin{aligned}\bar{r} &\leq 1 - b_\ell + \frac{\lambda}{1-\lambda}(b_h - b_\ell) + \frac{-\beta V_p'(b_\ell + \bar{r})}{(1-\beta)(1-\lambda)\theta^2 D'(\bar{r}\theta)} \\ \bar{r} &\geq 1 - b_\ell + \frac{\lambda}{1-\lambda}(b_h - b_\ell) + \frac{\beta V_p'(b_\ell + \bar{r})(b_h - b_\ell)}{(1-\beta)(1-\lambda)\theta\bar{q}}\end{aligned}$$

We then notice that the FOC implies that

$$V_p'(b_\ell + \bar{r})\left(1 + \frac{(b_h - b_\ell)\theta D'(\bar{r}\theta)}{\bar{q}}\right) \geq 0. \tag{14}$$

If $1 + \frac{(b_h - b_\ell)\theta D'(\bar{r}\theta)}{\bar{q}} < 0$ then increasing r with IC_h binding would yield $\bar{s} = b_h - (b_h - b_\ell)\frac{D(r\theta)}{\bar{q}} < w = b_\ell + r$, while decreasing r would yield $w = \bar{s} = b_h - (b_h - b_\ell)\frac{D(r\theta)}{\bar{q}} < b_\ell + r$. The FOC condition for r are then

$$\begin{aligned}\bar{r} &\geq 1 - b_\ell + \frac{\lambda}{1-\lambda}(b_h - b_\ell) + \frac{-\beta V_p'(b_\ell + \bar{r})}{(1-\beta)(1-\lambda)\theta^2 D'(\bar{r}\theta)} \\ \bar{r} &\leq 1 - b_\ell + \frac{\lambda}{1-\lambda}(b_h - b_\ell) + \frac{\beta V_p'(b_\ell + \bar{r})(b_h - b_\ell)}{(1-\beta)(1-\lambda)\theta\bar{q}}\end{aligned}$$

We then notice that the FOC implies that

$$V_p'(b_\ell + \bar{r})\left(1 + \frac{(b_h - b_\ell)\theta D'(\bar{r}\theta)}{\bar{q}}\right) \leq 0. \tag{15}$$

We conclude from this that it must be the case that when $\bar{s} = b_\ell + r$ we have

$$V_p'(b_\ell + \bar{r}) \geq 0 \Longleftrightarrow b_\ell + \bar{r} = \bar{s} < w_p^*.$$

We also notice that in this case we have

$$\bar{r} \geq 1 - b_\ell + \frac{\lambda}{1-\lambda}(b_h - b_\ell).$$

4) The monotonic case Which of the three types of solution is optimal is not easy to assess in general. But if we assume that $r - (b_h - b_\ell)(1 - D(r\theta)/\bar{q})$ is monotonic, thing are simpler. This is the case in particular if D is linear.

Suppose that $1 + (b_h - b_\ell)\theta D'(\bar{r}\theta)/\bar{q} > 0$; then we have $b_h - (b_h - b_\ell) D(r\theta)/\bar{q} < b_\ell + r$.

It follows that the optimal consumer prices with sponsored data and no exclusion are given by case 2.

Suppose that $1+(b_h - b_\ell)\,\theta D'(\bar{r}\theta)/\bar{q} < 0$ for all r; then we have $b_h - (b_h - b_\ell)\,D(r\theta)/\bar{q} > b_\ell + r$. In this case the optimal consumer prices with sponsored data and no exclusion are given by case 1.

B.2.2 Exclusion of the LB content

The last possibility is that $s > b_\ell$ and $\bar{s} > b_\ell$ so that the LB content is excluded.

A first point is that if $w_p^* \geq b_h$, it is not optimal to exclude the LB content when sponsored data is offered as an active. option. Indeed suppose first that $w = s + r < \bar{s} \leq b_h$. Then it would be optimal to increases $s + r$ until $s + r = \bar{s}$. But at this point the sponsored option is useless and the network can rely on uniform pricing. Suppose then $s + r > \bar{s} = w > s$ and that the paid content providers choose to sponsor and the HB content providers choose not to sponsor. Then we must have $s = b_h > 1$ and thus it is optimal to reduce r until $s + r = \bar{s}$, at which point again the option is useless.

Consider then the case $w_p^* < b_h$. In this case the only possibility with exclusion of the LB content is $s + r = w_p^* < \bar{s} = b_h - (b_h - s)\,\frac{D(r\theta)}{\bar{q}} > w_p^*$ and $s > b_\ell$ solves

$$\min_{s,r}\,(b_h - s)\,D(r\theta) \text{ st } s + r = w_p^*$$

which gives $s > b_\ell$ if

$$-D\left(\left(w_p^* - b_\ell\right)\theta\right) - (b_h - b_\ell)\,\theta D'\left(\left(w_p^* - b_\ell\right)\theta\right) > 0.$$

Such a solution can only arise if β is large and λ is large so that the gain on the paid content offsets the loss of the LB content.

B.3 Summary for the case $w_p^* > b_h$.

While the above characterization is quite complex, it becomes easier to understand when $w_p^* > b_h$. In this case the price charged to the paid content providers is always too low from the network perspective. As *(i)* the network is constrained on the content of the market if it wants to keep some free content on board and *(ii)* it is always optimal for the network to accommodate the HB content (the network can set $s = b_h$ and $r = w_p^* - b_h$ and do better than with exclusion of all free content), the presence of paid content will lead the network to inflate the consumer price of data allowance. This price inflation is the key difference with the case with only free content. To summarize this case we have:

Proposition 10 *Assume $w_p^* \geq b_h$. Then:*

(i) Under uniform pricing there is a threshold $\hat{\lambda}^u$ such that if $\lambda < \hat{\lambda}^u$, the network charges $s = b_\ell$ and $r > 1 - b_\ell$, whereas if $\lambda > \hat{\lambda}^u$ it charges $s = b_h$ and $r \geq 0$.

(ii) When sponsored data is offered the network sets

$$s = b_\ell;\ r > 1 + \frac{\lambda}{1-\lambda}(b_h - b_\ell) \text{ and } \bar{s} = b_h - (b_h - b_\ell)\frac{D(\theta r)}{\bar{q}}.$$

Proof. The point (i) follows from $\beta = 1$. The point (ii) follows from the fact that $V_p'(w) > 0$ in all the solutions derived above. ■

An immediate consequence is that the presence of paid content leads to more eviction of the LB content provider in the uniform case. Indeed exclusion becomes relatively more attractive as it allows raising the total price for paid content. More generally the consumption of LB content declines in all regimes.

Corollary 3 *Assume $w_p^* \geq b_h$. Then a ban of sponsored pricing raises exclusion. When exclusion doesn't occur, this ban reduces the consumer price of data without affecting the content price of data if either β is small or $1 + (b_h - b_\ell)\theta D'(\bar{r}\theta)/(1-\lambda)\bar{q}$ is negative. But the ban raises the consumer price of data if λ is small and $1 + (b_h - b_\ell)\theta D'(\bar{r}\theta)/\bar{q}$ is positive.*

Proof. The first point follows from the fact that we have shown that when sponsored data is proposed by the network it cannot be optimal to exclude the LB content. Thus exclusion can only occur with uniform pricing.

Then if there is no exclusion we always have $s = b_\ell$. We have under uniform pricing

$$r^{uniform} = 1 - b_\ell + \frac{\beta V_p'(r + b_\ell)}{(1-\beta)\theta}\left(\frac{-1}{\theta D'(r\theta)}\right)$$

while with sponsored data

$$r^{sponsored} \geq 1 - b_\ell + \frac{\lambda}{1-\lambda}(b_h - b_\ell) + \frac{\beta V_p'(r + b_\ell)}{(1-\beta)\theta}\frac{\min\left(\frac{-1}{\theta D'(r\theta)}, \frac{(b_h - b_\ell)}{\bar{q}}\right)}{1-\lambda}$$

We clearly have $r^{uniform} < r^{sponsored}$ if β is small. This is also the case if $1 + \frac{(b_h - b_\ell)\theta D'(\bar{r}\theta)}{(1-\lambda)\bar{q}} < 0$ because in this case $\frac{-1}{\theta D'(r\theta)} \leq \min\left(\frac{-1}{\theta D'(r\theta)}, \frac{(b_h - b_\ell)}{\bar{q}}\right)/(1-\lambda)$.

The reverse holds when λ is small and $\frac{-1}{\theta D'(r\theta)} > \frac{(b_h - b_\ell)}{\bar{q}}$. ■

Thus under the condition of the corollary we reach the same conclusion when there is little paid content and when there is only free content. A ban of sponsored data would raise the welfare for the LB content unless it triggers exclusion of this content. But the risk is that the ban induces more exclusion of content. An additional effect is that in the case where the paid content is not sponsored and the LB content is not excluded, the lower consumer price would also benefit paid content providers.

When the paid content providers choose sponsoring ($w = \bar{s}$) and the proportion of HB content is small, a difference arises with the case of free content. The reason for this is that the network would like to raise the price $\bar{s}$ charged to the paid content; since $\partial w/\partial r = \partial \bar{s}/\partial r < 1$, this is achieved by increasing the consumer price r even further than under uniform pricing (where $\partial w/\partial r = 1$).