Congestion Pricing and Net Neutrality*

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

Very preliminary and incomplete, June 2012

Abstract

We consider a network that intermediates traffic between content producers and consumers. The content is heterogenous in the cost of traffic. While, consumers do not know the traffic cost when deciding on consumption, a content producer knows his cost but may not control the consumption. The network observes only the resulting total cost of traffic and can charge a congestion price to one or both of the parties, along with an ex-ante hook-up fee to consumers.

We first show that, if the content is a paid content, the network charges only the content producers and capping congestion prices for content in this case is sub-optimal. In the case of free content, the network extracts some rent from content with congestion prices and may exclude some content. We show that there is efficient or excessive exclusion of traffic.

We then endogenize the choice of business model by allowing the content producers to choose between a paid model and a free model.

*We gratefully acknowledge France Télécom, in particular Marc Lebourges, for its intellectual and financial support.

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

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

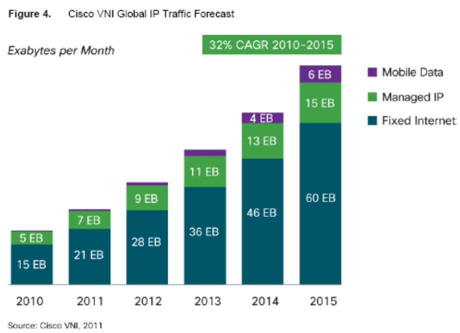
In this case, the network charges higher congestion prices to content but the cost is smaller as some content can stay under a paid model that would be excluded otherwise.

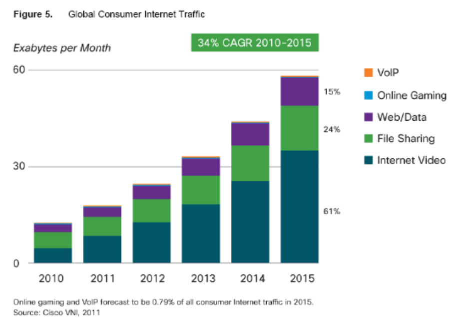
At last, we characterize an optimal mechanism which consists in letting the content producers choose between different public categories associated with different congestion prices for content and for consumers.

1 Introduction

Net neutrality is the object of an intense debate, with contrasted views on the way the operators of the physical network should treat various contents and on the relationship between content owners and Internet service providers. This debate involves many aspects, including issues of freedom and democracy. The economic part of the debate has focused mostly on two broad questions: i) whether the need for traffic management justifies that ISPs charge the suppliers of content on Internet a fee related to the volume of traffic or other cost dimensions; ii) whether a multi-tier quality structure with paid access to high quality delivery services is desirable.

The first question arises from the anticipation of congestion on the network and the need for ISPs to invest in the capacity so as to adapt to new exploding demand of traffic. The congestion issue is already present on the mobile data services due to last mile congestion, but traffic predictions (see the graph below) suggest that it will also arise in the fixed network. Moreover new services such as HD TV require massive investment in fiber technologies.





The second question arises from the difference in the technical requirement of various services. Some services require a high level of quality that a single layer of quality may not guarantee. The table below shows the need of various services in terms of three criteria for Quality of Service: latency which refers to the delay between origination and reception, jitter which refers to fluctuations and bandwidth.

Applications	Latence	Jigue	Bande passante
Mail	+	+	+
Partage de fichiers	+	+	++
Navigation Web	++	+	++
Jeux en ligne	+++	++	+
Vidéo à la demande	+	+++	+++
Appel en VoIP	+++	+++	+
Vidéo-conférence	+++	+++	+++

+ : sensibilité faible   +++ : sensibilité importante

Source: ARCEP

As technological convergence forces actors from very different traditions to work together (telecommunication, Internet, media), the question of pricing of Internet have generated a hot debate and regulatory activity (FCC release new rules in 2011,[1] EU is undertaking a consultation on the subject). The intensity of the debate reflects opposing views concerning the impact of price based traffic management on the various actors of the value chain. On the one hand there are concerns that pricing schemes may excessively crowd out content, reduce entry and content innovation, and jeopardize the traditional ecosystem that underlaid the success of Internet. On the other hand, one may fear that the lack of prices for content may result in insufficient investment and inefficient usage of the network.

The economic literature has addressed the issue from two perspectives, often combined. The first is the two-sided market model that allows to incorporate in the analysis of the pricing problem the two sides participating to Internet, consumers and content producers, and to discuss the implications of various price restrictions for the supply of content and the consumer surplus. The second perspective is the role that price discrimination may have in situations involving some hold-up problems, which allows to discuss the risk of ex-post expropriation of some content innovators.

The literature has highlighted interesting aspects of the economics of Net Neutrality, although it has not created a general consensus. One point that emerges from two-sided market models is that laissez-faire will not result in efficient pricing. This conclusion applies when there is market power but also when there is competition. Competition alleviates the problem but doesn't eliminate it. However it is quite difficult to reach a simple conclusion on the nature of public intervention that would restore efficiency. Allowing ISPs to charge a positive price for content may reduce the supply of content but it also intensifies competition to attract consumers, as increased customers participation can be leveraged with higher revenue on the content side of the market. Thus the precise nature of the intervention that would foster efficiency is unclear (see Economides and Tag (2009) for instance). One

[1] FCC, "Preserving the Open Internet:Final rule", http://www.gpo.gov/fdsys/pkg/FR-2011-09-23/pdf/2011-24259.pdf.

open question is whether addressing these inefficiencies is better done with regulation or with standard application of anti-trust policy by anti-trust authorities.

The analysis of a two-tier system has highlighted the effect of price-discrimination on content innovation. Prices allow ISPs to capture part of the innovation surplus which may reduce the incentives to innovate (Hermalin and Katz (2007), Grafenhofer (2011)), although the two-sided nature of the market alleviates the problem (Grafenhoffer (2011)). Moreover in models of congestion, a two-tier system implies a reduction of quality for the lower layer (Economides and Hermalin (2010)). These inefficiencies however must be compared with potential efficiencies arising from improving the quality for some services while maintaining the possibility of access to all contents.

The literature has also pointed to ambiguous effect of Net Neutrality on investment. Reduced ability to remunerate investment through adequate pricing strategy suggests that Net Neutrality may impede investment in capacity (Hermalin and Economides (2010), Bourreau, Kourandi and Valletti (2012)). However two-sided market effects may reverse the conclusion. As ISPs try to optimize the service offered to consumers, reduced traffic management possibility due to Net Neutrality along with increased traffic may lead to larger ISPs investment (Canon (2010)). This occurs if investment and traffic management are two substitutable instrument that ISPs can use to accommodate a traffic increase. Moreover, when there are two qualities of services, one of which being free, the risk is that the ISPs try to enhance the return on high quality by creating scarcity of bandwidth or degrading the quality for the free service. This calls for imposing minimal standards for QoS on basis services (Choi and Kim (2010), Bourreau, Kourandi and Valletti (2012)).

While the literature provided new insights, some specificity of Internet are not well discussed so far and deserves specific attention. In what follows we wish to highlight and discuss three aspects all related to the ubiquity of free services on Internet.

A. The missing price

One of the obvious specificity of Internet - and a key element in its success - is the ability of agents, producers but also consumers-producers, to innovate in terms of information services and business models. In particular, large successes on Internet have involved business strategies based on services offered for free to consumers. This was made possible by the relatively small variable cost of information goods and the potential for leveraging the information acquired on consumers by offering advertising-type services to third-parties (obvious examples are Google, Yahoo News or Facebook).

As a result of these new business models, we may view the economics of Internet as a case involving missing prices. Consumers don't pay for the content which prevent prices to play their role of signal of scarcity to consumers and of value to producers.[2]

From this perspective we may view the net Neutrality problem as follows. Efficient traffic management must ensure that various actors on the web internalize the cost they impose on the ecosystem, being a physical cost or a congestion cost. The question is then how to price congestion cost. In a network, the cost when communication takes place is the result of the interaction between two agents. In our case it is the consumer who receives the traffic and the content producer who sends the traffic. How the consumption is transformed into costs depends on factors that are usually known and controlled by the content producer or the network. Hence consumers can barely foresee the cost they impose on the network at the time they choose consumption. Thus some signal must inform them.

The answer to this problem in a market economy is that the cost is charged to the producer who imbeds it into the price he charges to consumers. In what follows we will make this transparent and show that the logic applies for a for-profit network only under some demand conditions. However, in situations where goods are offered free of charge to consumers this mechanism cannot

[2]The definition of two-sided markets proposed by Rochet and Tirole emphasizes that two-sidedness follows from some restriction on the set of prices available.

be used and alternative ways to signal costs must be invented. We will argue in the paper that the missing price perspective suggests that charging both sides is optimal, although whether content is charged a positive price or not depends on the context.

B. The choice of business model

The second aspect that deserves some attention is that decisions concerning the pricing of traffic may significantly impact the choice of business model by content producers. Indeed the business model adopted by content suppliers is endogenous. Typically suppliers will adopt the business model that maximizes their profit (or other objective if they are a not maximizing profit) and their survival chance faced with competition. To some extent, we may view the success of the free service model as resulting from the very low functioning cost on Internet. However which business model prevails depends on the market condition. In particular input prices affects business models in various ways. Clearly the free service model is the most affected by an increase in input prices as the producer has no way to share the cost with consumers. As we will argue, raising prices for content in the context of free services should trigger not only exit of some content but also a shift from free services to paid services.

C. Social vs. private value

While we emphasized that free services imply that consumers do not receive the right signal about the cost generated by their activity, we must realize that producers also face the wrong signal. Indeed content prices and profits usually signal to producers whether their product is valuable or not, which helps align the private incentives to invest and innovate with the social incentives. With a free service, the only signal that a producer obtains from the market is the consumption at zero price. Hence the signal is less flexible and of lower quality. This raises in particular the risk of supply overload as suppliers with large demand but little social value may flood the market. Notice that, concerning information goods, such an overload will have a cost in terms of traffic but also by diverting the limited cognitive resources of individuals from more efficient usage. As pointed by H. Simon (see also Anderson and de Palma (2009), van Zandt (2004), Falkinger (2007), (2008)), attention is a scarce resource and excessive supply of information may create information congestion. From this perspective one must evaluate the impact of various pricing schemes on the incentive to supply information.

Our paper

In this paper we consider a network that intermediates the traffic between content producers and consumers. The content is heterogenous in the cost of traffic (referred to as the load). Consumers know their preferences but not the load generated by their consumption; a content producer knows his load but has no direct control on the consumption. Content producers may receive income proportional to traffic, such as advertising revenue or direct utility for the producer, and may or may not charge a retail price for content. In the first part of the paper we take the business model as given; we endogenize it in a second part.

The network observes only the cost of traffic but not the consumption nor the load. Based on observed cost of traffic the network can charge a congestion price to one or both of the parties involved in traffic generation. Our question is then to determine the optimal price structure. In this context, the question of Net Neutrality amounts to ask whether the price charged to content producers should be regulated and at which level.

We assume that the market for content is competitive and that the network can charge a hook-up fee to consumers. We compare the case where content is sold at a competitive price and the case where content if free. If the content is a paid content, both the socially optimal and network tariff charge only the content producers for traffic with a price equal to cost. The cost is then fully incorporated into the retail price charged by the content producer to consumers.

In the case of free content, there are two types of inefficiencies. Firstly, consumers do not know the cost when they choose consumption. As a result their consumption may only reflect the average cost of traffic instead of the true cost. Hence, when charged the average cost of the network, there is overconsumption of the most traffic intensive content and under-consumption of the least traffic intensive content. Secondly, charging a price to content may exclude high load content as the producers cannot pass-on the cost to consumers when the good is free. The socially optimal tariffs are such that the consumers are charged a congestion price equal to the expected cost of traffic net of the value generated for content producers, which corresponds to a typical two-sided market price formula.

The network charges a positive price to content which shifts the rent toward him and reduces the opportunity cost of traffic. The content producers may be charged an exclusionary price (that prevents high load content to enter) if the load dispersion is large. A positive price for content allows the network to raise the value offered to consumers and thus profit, by reducing the congestion prices for consumers. When there is no exclusion the resulting price structure is constrained optimal, but the network may induce excessive exclusion. Net Neutrality can raise welfare only if the proportion of high cost to low cost content is neither too high nor too low.

We then endogenize the choice of business model. For this we assume that in order to collect a price for content, the content producer must put into place a costly micro-payment technology. The content producer will then choose between a high cost paid service or a low cost free service (where the only source of revenue is advertising). When the supply of content is competitive, competition will lead the producers to offer the best service to consumers subject to zero profit. We show that this implies that competitive producers will offer a free service provided that it is sustainable, thus provided that advertising revenues are sufficient to cover the expected cost. Therefore, increasing the congestion price charged to content producers may result into the adoption of a paid business model, with the associated increase in cost due to micro-payment. We then investigate the consequences and show that, despite inefficiencies, this reduces the social cost of congestion pricing because content that would be excluded if it was forced to stay free can now change its business model and survive. In this context no simple regulation unambiguously dominates the others.

We then characterize a mechanism designed to achieve transmission of the right signal to consumers. It involves screening different types of contents by a menu of tariffs and making the tariffs transparent to consumers, referred to as "category pricing". This requires that communication takes place prior to consumption. We show that the optimal allocation can be implemented by a offering a menu of categories associated with pairs of congestion prices (for receiver and sender). Each content producer chooses a category, and consumers are informed of the category prior to consumption. The price paid by consumers decreases with the price paid by the content. Faced to the menu, each content producer must trade-off the volume of consumption with the cost of traffic. The lowest load content will then opt for the highest consumption while the highest load content will opt for the lowest congestion cost.

2 Model

We analyze the tariff charged for traffic by a network (in the case of Internet, an Internet Service Provider (ISP)) to two sides of the market: consumers and content producers (or CPs). The mean demand for the content when consumers face a price p per unit of content is $E(q) = D(p)$. We assume that the consumption $D(0)$ of free goods is positive and bounded, and that there exists $\bar{p}$ such that $D(p) = 0$ for $p > \bar{p}$. We denote by $U(q)$ the representative consumer concave utility function defined by $U'(q) = D^{-1}(q)$ the inverse demand curve. Then the indirect utility function is defined by $S(p) \equiv \max_q U(q) - pq$.

We wish to emphasize the different impact of the traffic pricing depending on whether the content is free or paid. Of course, it is the content producer who decides on charging a price or not. To capture the fact that changing the tariff may change the business model we build a simple framework that allows for endogenous business model. Suppose that each unit of generate a net benefit a per unit for the content provider. This benefit, which can be positive or negative, includes the advertising revenue and other benefits of the CP but also the cost distributing the content. A content producer can choose to offer the good for free, which can be profitable if $a > 0$, or to charge a positive unit price p. But in order to charge a price, some micro-payment technology must be used that implies a unit cost $\mu \geq 0$ per unit of consumption.[3] Except in the pay content section where it will matter, we assume that a and μ are known and the same for all content, as this does not alter the main messages but simplify the formulas. We discuss the general case where all parameters are heterogenous in appendix. Because we wish to focus on the traffic management issue, we rule out monopoly or oligopoly distortions on the market for content and assume that each content is competitive with a continuum of suppliers.

Any transaction between CP and consumer generates a load for the network. More precisely, for any unit of consumption, each CP will generate a cost β (referred to as the load) to the network so the consumption of q units[4] of content with a load β generates a cost βq to the network. Each content is thus characterized by the cost it imposes on the network when a consumer download it. Admittedly we do not model explicitly congestion. One view is that the network needs to expand resources to maintain the quality of service and that β reflects this need.[5] In a more general set-up with explicit congestion, β would be interpreted as the shadow cost of congestion.

We assume that β is unknown to consumers and the network, but known to the content producers. More precisely, β can take on two values: $\beta \in$

[3]The analysis would be similar for a cost proportional to sales. We assume also that μ is paid even if a zero price is charged ex-post.

[4]For example, q may be the number of songs downloaded by the consumer while β is the bandwidth taken by each video. Alternatively one may view q as a number of subscriptions and β the traffic for one subscription.

[5]When the cost is only related to congestion, one may view βq as a cost that the network will bear ex-post to maintain the traffic.

$\{\underline{\beta}, \overline{\beta}\}$ with probabilities λ and $1-\lambda$, where $0 \leq \underline{\beta} < \overline{\beta}$. We denote by β^e the mean values of β. We refer to $\bar{\beta}$ as the high load content or HL content, and to $\underline{\beta}$ as the low load content or LL content.

While the content producer has information about the load, the level of consumption is determined by consumers. Moreover, the network cannot monitor β and q but observes the ex-post realization of cost βq and can charge any side for this cost. We restrict attention to linear traffic prices in the main part of the paper, $s\beta q$ to the content producers and $r\beta q$ to the consumers. Many networks (and in particular ISPs) charge hook-up fees to consumers in which case the network objective may internalize the surplus of consumers. The extent to which the hook-up fee allows the network to capture an increase in consumer surplus that traffic management generates depends on various factors and in particular the elasticity of participation to the network. We develop the case where the network maximizes the sum of its profit Π and consumer expected surplus SC.

We denote $V = SC + \Pi$ and refer to it as the network value. Maximization of $SC + \Pi$ occurs when consumers do not have private information about their expected surplus before joining the network. Then it is optimal for the network to maximize the joint expected surplus with consumers and to use the hook-up fee to share this surplus with the consumers. This is valid both for monopoly and competition, provided that consumers uses only one ISP while content may be distributed on all ISPs (thus we rule out exclusive contract between ISP and content). The market power of the network then determines the level of hook-up fee, but not the tariffs for traffic. [6]

In line with the current debate, we rule out hook-up fees for content and

[6] Maximization of Π would correspond to situation with zero or exogenous hook-up fees, and to situations where the marginal consumer for the participation decision does not consume the content. A simple model that illustate this outcome is the following. Participation utility (gross of transfers) is $v + \theta u(q/\theta)$, where v is fixed and θ is uniformly distributed on $(0, \bar{\theta})$. If the parameter θ is unknown to the consumer before he decides to participate, then the network maximizes $V = \Pi + SC$. If the parameter θ is known before the participation decision of the consumer and v is large, then the optimal hook-up fee is v and the monopoly network sets traffic prices to maximizes Π.

focus on the traffic sensitive price for content.

Timing

1. The network chooses the prices r and s.
2. CPs observes β and decide to exit, be free or paid. In the latter case, the price p is set at the competitive level.
3. Consumption takes place, as well as payments from consumers to CPs.
4. Traffic is observed, payments to the network take place.

Competition will lead content producers to maximize the value offered to consumers provided their profits are non-negative. Whenever $a \geq s\beta$, the competitive allocation corresponds to a free-content. Indeed a free-content cannot be displaced by a paid content and is profitable. Whenever $a < s\beta$, a free content allocation is not sustainable hence the only possibility for a competitive equilibrium is a paid content with $p = s\beta - a + \mu$. We assume that entry is always efficient if priced at the true marginal cost, i.e., $\bar{\beta} + \mu - a < \bar{p}$. At last, we denote the maximal congestion fee for free content for both types of CP as

$$\underline{s} = a/\underline{\beta} \text{ and } \overline{s} = a/\overline{\beta}, \text{ with } \underline{s} > \overline{s}.$$

Our objective in what follows is to discuss the determination of the prices r and s, and the impact of various regulations of s both on the business model and on the signal perceived by the consumer on the impact of their consumption. For this purpose we will start with simple benchmark cases, showing in particular that the first best easily achieved with paid content. Then, we will see how the presence of free content modifies this result and alter the way consumers are signaled the cost of their consumption on the network. This leads us to discuss the impact of various price regulation on the efficiency and the choice of business model made by the CPs.

Note at last that when evaluating welfare we will assume that content producers capture the full surplus of the advertisers so that a is also the social value of adds.[7]

3 Benchmark

3.1 Full information

To illustrate the model, consider the socially optimal prices for the case of a single content with full information. If negative prices for traffic could be implemented, the optimal allocation would implement free goods and uses the traffic prices to control for consumption with:

$$r = 1 - a/\beta \text{ and } s \leq a/\beta.$$

The tariff induces efficient consumption (consumers face a price $\beta - a$), while avoiding the micro-payment which is optimal because the network bears no transaction cost.[8]

The difficulty with the above solution is that it involves negative prices which may not be feasible. Suppose that negative prices cannot be used, then charging a non-negative price to CPs for the traffic may force the content producers to charge a positive retail price. In this context, we obtain:

Lemma 1 *Under full information, the socially (constrained) optimal allocation is obtained by :*
i) charging $r + s = 1$ *if* $a < 0$ *(paid content)*
ii) charging $r = \max(1 - a/\beta, 0)$ *and* $s \leq a/\beta$ *if* $a \geq 0$ *(free content).*

[7] An alternative is to consider that advertising induces some socially wasteful expenses in which case the weight attached to the website advertising revenue is less than one.

[8] This is a reasonable assumption given that the network can incorporate these payments into the overall bill for the service.

Proof. We solve

$$
\begin{aligned}
W &= \max_{s\geq 0, r\geq 0} U(q) - (\beta + \mu I_{p>0} - a)\, q \\
st\ U'(q) &= r\beta + p \\
p &= s\beta - a + \mu \ if\ s\beta - a > 0 \\
p &= 0\ if\ s\beta - a \leq 0
\end{aligned}
$$

A paid content is necessary for the participation of content if $a < 0$. Then we have $p = s\beta - a + \mu$ so that $U'(q) = (r + s)\beta - a + \mu$ and welfare is maximized at $r + s = 1$.

If $a \geq 0$, a paid content allocation is dominated by a free content with the same cost supported by consumers (with r' such that $r'\beta = r\beta + p$). Hence $s\beta - a \leq 0$, welfare is then maximal when $U'(q) = r\beta = \beta - a\beta$ if it is positive and $r = 0$ if it is negative. ■

Notice that setting $s = a/\beta$ in the case ii) of the proposition generates an allocation that is (constrained) efficient and such that the content producers receive zero surplus. The value $V = SC + \Pi$ is then equal to the maximal welfare W. This implies that the network maximizing V would implement the social optimum under full information about β. Thus our model is such that under full information a laissez-faire policy is optimal.

Provided that the net revenues are positive, the optimal allocation is obtained when the network intermediates the relationship between the content producers and the consumers so that the content is free and the consumer pays only the network. This corresponds to some form of vertical integration or a buy-and-resell model. Although this is not the topics of the current paper, notice that this is a solution adopted for some services by ISPs (as TV, VoD or phone services). This highlights an important property of our model, which is that introducing price for content is costly and thus that it should be avoided if alternative solutions can achieve the same consumption.

3.2 Paid content

Let us start by assuming that $a < 0$ so that all content is paid. To simplify and without loss of generality we normalize $\mu = 0$.

Under our assumption of a competitive market, the retail price for paid content is $p = s\beta$. Consumers on the other side face a price p for the content and $r\beta$ for the traffic. When deciding how much to consume facing $r > 0$, they must form some expectation over the traffic they will generate. For this, they may rely on the price of the content. Indeed if $s > 0$, the price of the content reveals some information about traffic: more expensive contents on average generate more traffic. For instance, a rational expectation equilibrium is defined as an allocation such that:

$$p = s\beta \tag{1}$$

$$q = D\left(rE\left(\beta \mid p\right) + p\right). \tag{2}$$

Equations (1) and (2) capture the idea that consumers will eventually realize that they tend to have more traffic when they consume more expensive content.

Remark that when all content is paid content, the profit of the content producers is zero. Therefore total welfare is equal to the sum of the consumer surplus and of the network's profit, i.e. $W = V$. Hence the network maximizes total welfare. Notice then that by setting $s = 1$ and $r = 0$, the network induces a retail price $p = \beta$ and obtains $V = E_{\beta}\left\{S\left(\beta\right)\right\}$ which is the maximal total welfare that can be generated. Hence this must be optimal.

Proposition 1 *With only paid content, the network maximizes total welfare with optimal tariffs* $s = 1$ *and* $r = 0$*.*

Proof. The first-best is obtained for $t = s = 1$ and $r = 0$. ■

Hence social optimality in the case of paid content requires that the content producers pay for traffic. This is because this induces the best price signal for consumers as the content producers are in the best position to set this signal. In this case, the optimum has a very standard structure: the network sells the traffic at cost to the content producer who then set the retail price.

Note that provided that $s > 0$, we have $E(\beta \mid p) = \beta$. It follows that an alternative way to implement the first-best is $0 < s < 1$ and $r = 1 - s$. However it cannot be the case that $s = 0$ because with no price s, the retail price of content p is not informative. Hence the robust conclusion is that there is a positive price for content.

The solution with $r > 0$ and $s > 0$ is more complex at it requires 2 prices (it cannot be the case that $s = 0$) and would be dominated if there were some noise in the consumers' inference process. To see that suppose that a is random with support $[\underline{a}, \bar{a}]$ and unknown to consumers. The quality of the consumers' information depends on the precision of the signal transmitted by prices on the load. This suggests that $s = 1$ is optimal. Indeed, given the above inference by consumers the value can be written as

$$V = SC + \Pi = E\left\{S\left(rE\left(\beta \mid p\right) + p\right) + \left(r + s - 1\right)E\left(\beta \mid p\right)D\left(rE\left(\beta \mid p\right) + p\right)\right\}.$$

If there is noise with $\bar{\beta} - \underline{\beta} < \bar{a} - \underline{a}$, then at any $r > 0$ there is a positive probability that $rE(\beta \mid p) + p \neq \beta$ and therefore $V < E_\beta\left\{S(\beta)\right\}$. On the other hand setting $s = 1$ and $r = 0$ still implements the first-best because with these prices the consumers' expectation over β are irrelevant and the retail price reflects the true cost.

Hence when all content charges retail prices, there is a strong argument for a traffic price equal to marginal cost paid by content producers.

4 Free content

Let us now turn to the case where all the content is free. In this case we assume that μ is very large so that a paid content is not an option. Notice that a free content is profitable only if $a > 0$ which we assume from now on.

Absent any price for traffic, all content producers are active but a positive price s may induce some exclusion. Notice that a price s charged to content is not reflected in an equivalent increase in the cost supported by consumers. This has two implications:

i) If the network wishes to reduce the consumption, it has to do so with a price $r > 0$ to consumers; the reduction is then uniform across contents.

ii) If the network wishes to reduce selectively the consumption of HL contents, he can only do so with a price s high enough that the HL content producers exit from the market.

Notice also that we are restricting ourselves to a single price r and a single price s that are determined ex-ante. In particular we rule out for the moment sophisticated schemes that tries to intermediate the relation between the CPs and the consumers. We will discuss these schemes in the last section of the paper.

A second important point is that, unlike paid content, the free content producers obtain a positive surplus that is not eliminated by competition. As we assume away wasteful dissipation of this surplus, this means that the objective of the network and social welfare do not coincide even if the network internalizes consumer surplus. Then the network will try to capture the CP surplus which creates an incentive to tax CPs.

As a price $s > 0$ allows to appropriate part of the surplus of free content producers, we expect the network to exclude too much content, but we will see that this point is not obvious.

Facing a price s, the CPs stay on the market if they anticipate a non-negative profit , hence if $s\beta < a$. This implies that the volume of free content is:

$$M = \begin{cases} 1 & \text{if } s \leq \bar{s} = a/\bar{\beta}; \\ \lambda & \text{if } \bar{s} < s \leq \underline{s} = a/\underline{\beta}. \end{cases}$$

The average traffic load is then $\phi = E[\beta \mid s]$ given by

$$\phi = \begin{cases} \beta^e & \text{if } s \leq \bar{s} \\ \underline{\beta} & \text{if } \bar{s} < s \leq \underline{s}. \end{cases}$$

Increasing the tariff s above $\bar{s}$ excludes HL content and thus reduces the average load.

In a rational expectation equilibrium, consumers correctly anticipate the mean load and the traffic cost $r\phi$. As they do not pay for content they will consume $D(r\phi)$ for each content.

We can then define the consumer surplus and the profit:

$$SC = M.S(r\phi), \ \Pi = M.(r+s-1)\phi D(r\phi).$$

As above, the network maximizes the joint surplus with consumers

$$V = M[S(r\phi) + (r+s-1)\phi D(r\phi)]$$

The term $S(r\phi) + (r+s-1)\phi D(r\phi)$ captures the incentive to maximize the joint surplus of the network and consumers for a given value of s (hence of ϕ). Internal efficiency is then achieved by setting a price

$$r = \max(1-s, 0). \tag{3}$$

When $s \leq 1$,we have $V = M.S((1-s)\phi)$ which is the consumer surplus from free content, accounting for the cost reduction that the tax on content generates. The network chooses s by comparing

$$\begin{aligned} V(\bar{s}) &= \max_{r \geq 0} [S(r\beta^e) + (r+\bar{s}-1)\beta^e D(r\beta^e)] \ \text{ if } s = \bar{s} \\ &\quad \text{and} \\ V(\underline{s}) &= \lambda \max_{r \geq 0} \left[S\left(r\underline{\beta}\right) + (r+\underline{s}-1)\underline{\beta} D\left(r\underline{\beta}\right)\right] \ \text{ if } s = \underline{s} \end{aligned}$$

Note with the low CP price ($s = \overline{s}$), there is no exclusion but some rent is left to the LL content and the cost of traffic is high while with the high price ($s = \underline{s}$), the HL content is excluded but all the LL content rent is captured by the network and the cost of traffic is low.

Lemma 2 *There exists* $\lambda^* > 0$ *such that the network excludes the HL content with a price* $\underline{s}$ *if* $\lambda > \lambda^*$.

Proof. Denoting $R = r\phi$ the expected congestion cost for consumers, there is exclusion if

$$\lambda \max_{R \geq 0} \left[S(R) + \left(R + (\underline{s} - 1) \underline{\beta} \right) D(R) \right] - \max_{R \geq 0} \left[S(R) + (R + (\bar{s} - 1) \beta^e) D(R) \right] > 0$$

Notice that there is no exclusion if $\lambda = 0$ (because demand is positive at price $(1 - \bar{s}) \beta^e$) and there is exclusion if $\lambda = 1$ (because $\underline{s} > \bar{s}$).

The slope is

$$\max_{R \geq 0} \left[S(R) + \left(R + (\underline{s} - 1) \underline{\beta} \right) D(R) \right] - (\bar{s} - 1) D(\max(1 - \bar{s}, 0) \beta^e) \left(\underline{\beta} - \bar{\beta} \right)$$

and the second derivative is 0 if $1 < \bar{s}$ and $D'((1 - \bar{s}) \beta^e) (1 - \bar{s})^2 \left(\underline{\beta} - \bar{\beta} \right)^2 < 0$ if $1 > \bar{s}$. The condition is concave or linear which implies that it hold for λ above a threshold λ^* ■

Typically the network will tax the content when the effect of exclusion on the average cost is large and when the average consumer surplus per unit of consumption is small.

At $\lambda = \lambda^*$ we have (denoting $R = r\phi$) :

$$\begin{aligned} \lambda^* \max_{R \geq 0} \left[S(R) + \left(R + a - \underline{\beta} \right) D(R) \right] &= \max_{R \geq 0} \left[S(R) + \left(R + \left(a/\bar{\beta} \right) \beta^{e*} - \beta^{e*} \right) D(R) \right] \\ \text{where} \quad \beta^{e*} &= \lambda^* \underline{\beta} + (1 - \lambda^*) \bar{\beta} \end{aligned}$$

This defines λ^* as a function of all parameters (including a). The condition can be written:

$$\begin{aligned}
& \max_{R\geq 0}\left[S\left(R\right)+\left(R+a-\underline{\beta}\right)D\left(R\right)\right] \\
= & \max_{R\geq 0}\left[S\left(R\right)+\left(R+a\underline{\beta}/\bar{\beta}-\underline{\beta}\right)D\left(R\right)+\frac{\left(1-\lambda^{*}\right)}{\lambda^{*}}\left(S\left(R\right)+\left(R+a-\bar{\beta}\right)D\left(R\right)\right)\right]
\end{aligned}$$

In particular we see that λ^{*} goes to 1 when a goes to 0.

When $a\geq\bar{\beta}$, we have $\bar{s}>1$ and the value of r is always zero so that we have

$$\begin{aligned}
S\left(0\right)+\left(a-\underline{\beta}\right)D\left(0\right) &= S\left(0\right)+\left(a\underline{\beta}/\bar{\beta}-\underline{\beta}\right)D\left(0\right) \\
&\quad +\frac{\left(1-\lambda^{*}\right)}{\lambda^{*}}\left(S\left(0\right)+\left(a-\bar{\beta}\right)D\left(0\right)\right) \\
\Longrightarrow &\ \lambda^{*}=\frac{S\left(0\right)+\left(a-\bar{\beta}\right)D\left(0\right)}{S\left(0\right)+\left(a-\bar{\beta}\right)D\left(0\right)+a\left(\frac{\bar{\beta}-\underline{\beta}}{\bar{\beta}}\right)D\left(0\right)}
\end{aligned}$$

which converges to $\bar{\lambda}=\frac{\bar{\beta}}{2\bar{\beta}-\underline{\beta}}$ when a goes to infinity. Notice that $1/2<\bar{\lambda}<1$.

Hence λ^{*} decreases from 1 to $\bar{\lambda}$ with a.

Unlike the paid-content case, the choice of the network is no longer optimal as it doesn't account for the surplus of the content. Indeed, using the definitions of $\underline{s}$ and $\bar{s}$, total welfare can be written as

$$\begin{aligned}
W\left(\bar{s}\right) &= V\left(\bar{s}\right)+\lambda(a-\underline{\beta})D\left(\bar{R}\right) \\
W\left(\underline{s}\right) &= V\left(\underline{s}\right)
\end{aligned}$$

where $\bar{R}=\beta^{e}\max\left(1-\bar{s},0\right)$. An immediate remark is that W coincides with the objective V when there is exclusion. We then find:

Proposition 2 *With only free content, there is excessive exclusion by the network.*

1. *If $a\geq\bar{\beta}$, the optimal level of s is $\bar{s}$, inducing $r=0$ and no exclusion.*

2. *If $\bar{\beta} > a \geq \underline{\beta}$, the optimal level of s is $\bar{s}$ with no exclusion or any $s \in [1, \underline{s}]$ with exclusion. Exclusion is optimal for λ large.*

3. *If $\underline{\beta} > a$, the optimal level of s is $\bar{s}$ with no exclusion or $s = \underline{s}$ with exclusion. Exclusion is optimal for λ large.*

Proof. See appendix ■

The network who internalizes the consumer surplus acts as a monopoly on behalf of the consumers (who pay then the cost ϕ but receives revenues s per unit of consumption). The revenue s is then redistributed to consumers in the most efficient way, that is by a reduction of r below ϕ along a fixed transfer if $s > 1$. In contrast with the case of paid content, the situation is now one with a two-sided market where the network extracts revenue on both sides. As a result there is excessive exclusion by the network when the content is free.

The optimal regulation is quite simple when $a \geq \bar{\beta}$. Indeed in this case imposing marginal cost pricing $s = 1$ leads to the optimal allocation. To see that notice that in this case lemma 1 shows that the full information optimum is at $r = 0$ and free content. But this is precisely what the network does if it is required to price s at cost.

Corollary 1 *If $a \geq \bar{\beta}$ an optimal regulation is marginal cost pricing $s = 1$.*

Proof. Immediate from above. ■

Of course this case is not the most interesting as the optimum is trivial. If $a < \bar{\beta}$ the HL content exits at $s = 1$ so that a regulation at $s = 1$ is clearly dominated by laissez-faire where the network may (efficiently) decide to let the HL content survive. Then an alternative is a Net Neutrality regulation that impose $s = 0$,which we examine now.

A first immediate remark from above is that, as for the paid content, a small positive price is preferable to a zero price provided that the price stays below $\bar{s}$. This is because the price r charged to consumers under Net

Neutrality is excessive ($r = 1$) and consumption is sub-optimal. Increasing s induces the network to reduce r which is welfare improving.

In particular, if the regulator knows that there is no risk of exclusion $(\lambda < \lambda^*)$*, then a zero price regulation is dominated by laissez-faire.*

The key question is whether Net Neutrality may dominate laissez-faire if $\lambda > \lambda^*$. The benefit of allowing exclusion under laissez-faire is two-fold:

i) the average cost of traffic is lower;

ii) the consumption is larger for the non-excluded content because r is low (due to the sea-saw effect that induces the network to reduce the price r when it can charge a positive price s).

On the other hand, a zero price regulation avoids exclusion of potentially valuable content. Whether such a regulation is optimal or not then depends on the social cost of exclusion.

When $s = 0$, the network sets a consumer price $r = 1$. Surplus is then

$$W(0) = S(\beta^e) + aD(\beta^e)$$

which is increasing in λ (as $a > 0$).

Let us assume that D is linear or convex. The optimality of a zero price s regulation depends on the sign of $W(\underline{s}) - W(0)$, which is concave in λ, negative when $\lambda = 0$ and positive when $\lambda = 1$ (as $W(\underline{s})$ is the maximal welfare when $\lambda = 1$ and $\underline{R} < \underline{\beta}$). *Hence the regulation is optimal below a threshold* λ_0.

To evaluate the policy we need to evaluate whether $\lambda_0 > \lambda^*$. Indeed if $\lambda_0 < \lambda^*$, there will be no value of λ for which the policy is optimal. We then have

Lemma 3 *Assume that D is linear or convex, then there exists λ_0 such that $Ws(0) > W(\underline{s})$ if and only if $\lambda < \lambda_0$. Moreover $\lambda_0 > \lambda^*$ if a is small, or if a is large and* $D\left(\frac{\bar{\beta}^2}{2\bar{\beta}-\underline{\beta}}\right) > \frac{\bar{\beta}}{2\bar{\beta}-\underline{\beta}}D(0)$.

Proof. See appendix. ■

The zero price regulation then dominates laissez-faire if and only if $\lambda_0 > \lambda^*$ *and* $\lambda \in (\lambda^*, \lambda_0)$.

5 Endogenous business model

In this section we extend the analysis by allowing the content producers to choose between a paid model and a free model. We are interested in understanding how the network can screen between traffic intensive and other contents, by inducing some content to change its business model. To focus on this issue, we make the following assumptions

Assumption 1 $\underline{\beta} < a < \bar{\beta}$.

The condition ensures that the LL content should be free of charge, while absent transaction costs the consumers should pay a positive price for the HL content. As we have seen above the optimal regulation if all content should be free ($a > \bar{\beta}$) is obvious and corresponds to marginal cost pricing $s = 1$. We discuss the choices of the network for this case at the end.

The choice of business model for the content depends on the price s. When $s = 0$ it is free, then as s increases more and more traffic intensive content becomes paid:

1. If $s \leq \bar{s}$, then all content is free;

2. If $\bar{s} < s \leq \underline{s}$, then the LL content is free, while the HL content is paid at price $p = s\bar{\beta} + \mu - a$;

3. If $\underline{s} < s$, then all content is paid.

Note that assumption 1 is equivalent to $\bar{s} < 1 < \underline{s}$.

In contrast to the previous section, the content that was excluded before is now sold at a positive price. Effective exclusion only occurs if s is so high that $p > \bar{p}$ (see below). Thus the social cost of exclusion is lower and we expect that laissez-faire is less detrimental. From the point of view of the network, the cost of raising s should also be lower because the consumer loss due to exclusion is lower.

Let us first derive a preliminary result.

Lemma 4 *For any price s, the platform optimal choice of consumer price is $r = \max(1-s, 0)$*

Proof. See the Appendix B on the general case. ■

As discussed in the appendix this conclusion is fairly general and depends only on the fact that the network maximizes the value $V = \Pi + SC$.

Let us consider now the optimal strategy of the network. First we notice that as long as $s \leq \bar{s}$, the analysis involves only free content and thus the previous section results apply. If the network decides to induce only free content it chooses

$$s = \bar{s} \text{ and } r = 1 - \bar{s}.$$

The value of the network is then

$$V^f = S\left(\left(1-\bar{s}\right)\beta^e\right).$$

For a higher price s, some content is paid. It is easy to see that because we have assumed that $\underline{s} > 1$, the network will never charge a price that induces only paid content. This is because for the LL content, the optimal consumer price is zero and the network can reap the full surplus with a price $s = \underline{s}$. Thus it is better to charge a large price below $\underline{s}$ than a price strictly above.

Lemma 5 *The network never charges a price s such that all content is paid.*

Proof. Suppose that $s > \underline{s}$. Then the price is $p = s\beta + \mu - a$ and the value is

$$E\left\{S\left(\left(r+s\right)\beta + \mu - a\right) + \left(r+s-1\right)\beta D\left(\left(r+s\right)\beta + \mu - a\right)\right\}$$

Because $\underline{s} > 1$, the optimum is $r = 0$ and $s = \underline{s} + \varepsilon$ with ε very small. The profit is then close to

$$\lambda \left[S(\mu) + (\underline{s} - 1)\underline{\beta} D(\mu)\right] + (1 - \lambda)\left[S\left(\underline{s}\bar{\beta} + \mu - a\right) + (\underline{s} - 1)\bar{\beta} D\left(\underline{s}\bar{\beta} + \mu - a\right)\right]$$

But then with a reduction slightly at $\underline{s}$ the network would induce the LL content to be free and obtain

$$\lambda \left[S(0) + (\underline{s} - 1)\underline{\beta} D(0)\right] + (1 - \lambda)\left[S\left(\underline{s}\bar{\beta} + \mu - a\right) + (\underline{s} - 1)\bar{\beta} D\left(\underline{s}\bar{\beta} + \mu - a\right)\right]$$

which is larger. Hence the LL content is always free. ■

Thus the alternative to all free content is a price that induces only the HL content to be paid. In this case, consumers can perfectly infer the load from the business model. The profit is then for $s > \bar{s}$:

$$\begin{aligned} V &= \lambda\left(S\left(r\underline{\beta}\right) + (r + s - 1)\underline{\beta} D\left(r\underline{\beta}\right)\right) \\ &\quad + (1 - \lambda)\left[S\left(r\bar{\beta} + s\bar{\beta} + \mu - a\right) + (r + s - 1)\bar{\beta} D\left(r\bar{\beta} + s\bar{\beta} + \mu - a\right)\right] \end{aligned}$$

It is quite immediate that

Lemma 6 *When the network chooses to induce some paid content, it sets prices $r = 0$ and $s^p \in (1, \underline{s}]$. There exists $\lambda^x < 1$ such the price s^p is increasing with $\lambda < \lambda^x$ with $s^p = 1$ at $\lambda = 0$ and $s^p = \underline{s}$ for $\lambda > \lambda^x$.*

Proof. See appendix ■

The optimal price for consumers is always $r = 0$ as the net revenue is positive. On the other hand the optimal price for content producers balances the total revenues generated by the two types of content. Indeed the revenue from the free content increases with s while the total revenue generated by the paid content (traffic plus consumer hook-up fee) is maximized when $s = 1$. Hence the optimal price s^p increases when the share of LL content increases.

From the above result the maximal possible price for content is $\underline{s}$. At this price we may distinguish two cases depending on whether the paid content is excluded from the market or just from the free segment. Define the cut-off

$$\hat{s} = \frac{\bar{p} + a - \mu}{\bar{\beta}},$$

corresponding to the maximal congestion price at which the HL paid content faces a positive demand.

- If $\underline{s} < \hat{s}$, then the HL content is never excluded from the market (this occurs when $\bar{\beta}/\underline{\beta} < 1 + \frac{\bar{p}-\mu}{a}$)
- If $\underline{s} > \hat{s}$, then the HL content may be excluded from the market.

In the latter case we can decompose the problem by defining

$$\begin{aligned} \hat{V}^p \quad = \quad & \max_{\bar{s} \leq s \leq \hat{s}} \lambda \left[S(0) + (s-1) \underline{\beta} D(0) \right] \\ & + (1-\lambda) \left[S\left(s\bar{\beta} + \mu - a \right) + (s-1) \bar{\beta} D\left(s\bar{\beta} + \mu - a \right) \right] \end{aligned} \quad (4)$$

We then have

$$V^p = \max \left(\hat{V}^p, \lambda \left[S(0) + (\underline{s} - 1) \underline{\beta} D(0) \right] \right)$$

and the HL content exits the market for $\lambda > \lambda^x$. Let us now turn to the choice of congestion price. The network compares V^f and V^p defined in (6). Given that V^p is larger or equal than the value with free content and exclusion, we expect that the network excludes more often the HL content from the free segment:

Proposition 3 *There exists λ^{**}, with $0 < \lambda^{**} \leq \lambda^*$ such that all content is free if and only if $\lambda \leq \lambda^{**}$.*

Proof. See appendix ■

The proposition shows that the network excludes at least as much when the business model is endogenous than when only free content is available. A sufficient condition for $\lambda > \lambda^{**}$ is

$$\mu < \frac{\lambda}{(1-\lambda)}\left(\bar{\beta} - a\right)\underline{\beta}/\bar{\beta}.$$

Indeed this condition ensures that $S\left((1-\lambda)\left(\bar{\beta}+\mu-a\right)\right) > V^{f}$. Using $V^{p} > \lambda S(0) + (1-\lambda) S\left(\bar{\beta}+\mu-a\right)$ and the convexity of S, we obtain $V^{p} > S\left((1-\lambda)\left(\bar{\beta}+\mu-a\right)\right)$ hence $V^{p} > V^{f}$.

Whether the HL content is paid or excluded for $\lambda > \lambda^{**}$ depends on the value $\underline{s}$ and μ.

Corollary 2 *When the business model is endogenous*

1. *If $\underline{s} < \hat{s}$ then $\lambda^{**} < \lambda^{*}$ and the HL content is paid for $\lambda > \lambda^{**}$.*
2. *When $\underline{s} > \hat{s}$,*
   - *if μ is not too large, then $\lambda^{**} < \lambda^{*}$ and the HL content is paid for $\lambda^{**} < \lambda < \lambda^{x}$ (where $\lambda^{*} < \lambda^{x}$);*
   - *otherwise $\lambda^{**} = \lambda^{*}$ and the HL content exits if $\lambda > \lambda^{**}$.*

Proof. If $\underline{s} < \hat{s}$ then $V^{p} > \lambda\left[S(0) + (\underline{s}-1)\underline{\beta}D(0)\right]$ implies that $V^{p} > V^{f}$ at λ^{*} and thus $\lambda^{**} < \lambda^{*}$. The same holds for $\underline{s} > \hat{s}$ if $\lambda^{x} > \lambda^{*}$. The only case where $\lambda^{**} = \lambda^{*}$ is when $\underline{s} > \hat{s}$ and $\lambda^{x} < \lambda^{*}$ which occurs if at λ^{*}, $\lambda^{*}\left[S(0) + (\underline{s}-1)\underline{\beta}D(0)\right] = V^{f} > \hat{V}^{p}$. This occurs if μ is large.

As $\bar{s} < 1$, we have

$$\hat{V}^{p} > \lambda^{*} S(0) + (1-\lambda^{*}) S\left(\bar{\beta}+\mu-a\right) > S\left(\left(\bar{\beta}-a\right)\frac{\beta^{e*}}{\bar{\beta}}\right) \; if \; \mu \text{ is small}$$

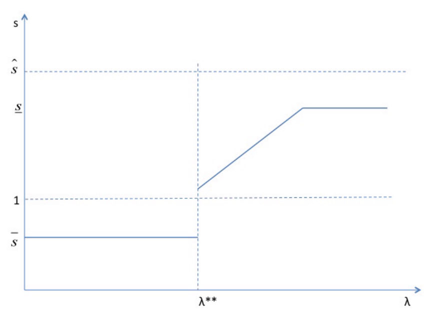
so that we have $\lambda^{x} > \lambda^{*}$ for μ small. ■

The corollary highlights the new features compared to the free-content model. The first statement says that when full extraction of the rent of the

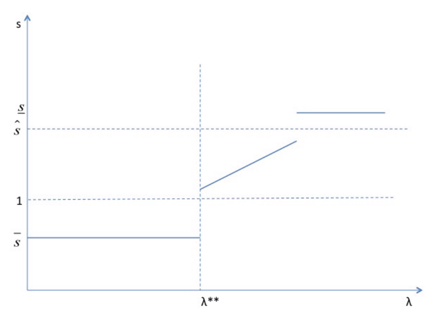
LL content is compatible with the survival of the HL content as paid content, this will occur more often then when it implies exit. The second statement derives conditions under which partial extraction of the rent of the LL content maintaining the participation of HL content is preferable to all free content or full extraction of the rent of the load content but with exit of the HL content.

The last statement says that if μ is large, then the paid content is never induced by the network policy. This is because this requires too high retail prices for the paid content which reduces the consumer surplus from paid content so that the network prefers to extract the full rent of the LL content.

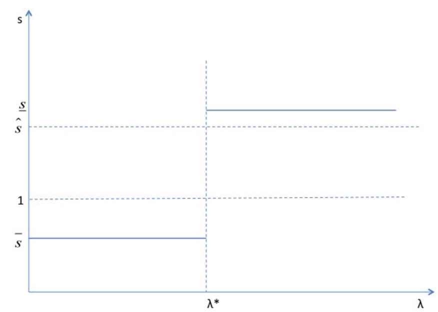
The possible configurations are illustrated in the next three figures.



Case with no exclusion



Case with exclusion



Case with only free content (large cost μ)

Let us now discuss the regulation of s. Notice that unlike the previous section, there is now some scope for regulation of prices at levels that differ from $\bar{s}$ and $\underline{s}$. First whenever $\underline{s} > \hat{s}$, reducing the price from $\underline{s}$ to below $\hat{s}$ when $\lambda > \lambda^x$ can only raise welfare. Indeed as $\hat{s} > 1$ the price r stays at zero so that the welfare is unchanged for the LL content but the HL content is present in the market while it was not before.

Moreover in the range where $\hat{s} \geq s \geq 1$, the HL content is paid and we have

$$W = \lambda \left[S(0) + \left(a - \underline{\beta}\right) D(0) \right] + (1-\lambda) \left[S\left(s\bar{\beta} + \mu - a\right) + \left(s\bar{\beta} - \bar{\beta}\right) D\left(s\bar{\beta} + \mu - a\right) \right] \quad (5)$$

which is maximized at $s = 1$. Thus the network sets the price s^p at an excessive level.

This suggests that a price cap may improve the situation. The difficulty with a price cap is that the network may charge a positive price r if the price-cap is too tight.

Proposition 4 *The optimal regulated price s^R is less or equal than* 1.

Proof. See Appendix B on the general case ■

This result is general, and does not depend on the particular distribution of β, a or μ. The difficulty with this result is that the precise value of s^R is complex to determine. An ex-ante regulation would require detail information about the all the parameters prior to observing the market. When this is the case the regulator may opt for simple rules such as a zero price s or cost orientation $s = 1$.

However while the result shows that some level of s^R would improve on laissez-faire, it is not the case that any level below 1 would do so. As we show now, there is no simple rule that is always optimal. For the discussion, we now focus on comparing the laisser-faire with three regulatory options: cost orientation ($s^R = 1$), a price-cap $s \leq 1$ and a zero price $s^R = 0$.

Note first that when $\lambda > \lambda^{**}$, a regulated price at $s = 1$ dominates laisser-faire. Indeed, imposing $s = 1$ induces the HL content to be paid with a price $p = \bar{\beta} + \mu - a$ that is efficient. This improves welfare because either the HL content was excluded from the market or the price p was not efficient. However imposing a fixed regulated price $s = 1$ may be suboptimal if $\lambda < \lambda^{**}$ because it induces the HL content to be paid even for high μ.

One may think that a price cap at $s = 1$ would solve this issue but unfortunately a price cap may be suboptimal if it induces the network to choose a low s with only free content while it would be socially optimal to have paid HL content even if this means laissez-faire. More generally let $V(s)$ be the value of the network at congestion price s (when it chooses r freely) and $W(s)$ be the total welfare. We then have

$$W(\bar{s}) - V(\bar{s}) = \lambda \left(a - \bar{s}\underline{\beta}\right) D\left((1 - \bar{s})\beta^e\right)$$

and for $s > \bar{s}$

$$W(s) - V(s) = \lambda \left(a - s\underline{\beta}\right) D(0).$$

These relations express that at $\bar{s}$ or above, the network captures all the profit of the HL content. Using that we see that

$$\begin{aligned} W(\bar{s}) - W(s) &> V(\bar{s}) - V(s) \\ if \ \left(a - \bar{s}\underline{\beta}\right) D\left((1-\bar{s})\beta^{e}\right) &> \left(a - s\underline{\beta}\right) D(0) \end{aligned}$$

A simple transformation shows that this holds when

$$s > s^{e} = \frac{D\left((1-\bar{s})\beta^{e}\right)}{D(0)}\bar{s} + \left(1 - \frac{D\left((1-\bar{s})\beta^{e}\right)}{D(0)}\right)\underline{s}$$

A price-cap at $s > \bar{s}$ is improving welfare if there is no excessive incentive for the network to induce free-content, hence if $s > s^{e}$.

Corollary 3 *If $s^{e} < 1$, then a price-cap at $s = 1$ improves welfare compared to laissez-faire.*

Proof. Suppose that $s^{e} < 1$ and a price-cap is imposed. Then welfare improves if $\lambda^{**} < \lambda$ and the network opts for $s = 1$, and welfare is unchanged if $\lambda < \lambda^{**}$. If $\lambda > \lambda^{**}$ the network opts for $s = \bar{s}$, then we have $V(\bar{s}) > V(1)$ which implies because $1 > s^{e}$ that $W(\bar{s}) > W(1) > W(s^{**})$ where s^{**} is the level under laissez-faire. ■

Notice that $D\left((1-\bar{s})\beta^{e}\right)/D(0)$ lies between $D\left((1-\bar{s})\bar{\beta}\right)/D(0)$ and $D\left((1-\bar{s})\underline{\beta}\right)/D(0)$, therefore s^{e} is always strictly between $\bar{s}$ and $\underline{s}$. In particular a sufficient condition for $s^{e} < 1$ is

$$\frac{D\left(\bar{\beta} - a\right)}{D(0)}\bar{s} + \left(1 - \frac{D\left(\bar{\beta} - a\right)}{D(0)}\right)\underline{s} < 1$$

which holds if $\underline{s}$ is not too large.

Let us now turn to a regulation at $s = 0$ (Net Neutrality). First notice that a price cap at cost may or may not be more desirable

Corollary 4 *If $s^{e} > 1$, then a price-cap at $s = 1$ improves welfare compared to a regulation at zero price for CPs.*

Proof. Suppose that $s^e > 1$ and a price-cap is imposed. If the network chooses $s = \bar{s}$, as $W(\bar{s}) > W(0)$, welfare improves compared to $s = 0$. When the network opts for $s = 1$ we have $V(1) > V(\bar{s})$ which implies because $1 < s^e$ that $W(1) > W(\bar{s}) > W(0)$. ■

Now suppose the only choice is between net neutrality and laissez-faire. Then the new feature compared to the section on free-content is that both the social and the private value of raising the congestion price for content is smaller because the HL content can choose to be paid. One consequence was that $\lambda^{**} < \lambda^{*}$. A second consequence is that the critical level of λ_{00} below which a regulation at $s^R = 0$ is preferable to letting the network set s above 1 is lower than λ_0. Hence it is not clear whether allowing the content to choose the business model makes the zero price regulation more or less attractive. But it reduces the potential cost of laissez-faire

Case $\bar{\beta} < a$. For completeness we discuss briefly the case where $\bar{\beta} < a$. If the network decides to induce only free content it chooses $s = \bar{s}$ and $r = 0$. The value of the network is then

$$V^f = S(0) + (\bar{s} - 1)\beta^e D(0).$$

As before the network never charges a price s such that all content is paid and sets $r = 0$ if some content is paid. The analysis of s^p is similar except that because $\bar{s} > 1$, we have $s^p = \bar{s}$ for μ small. Then the threshold λ^{**} is defined as for the case $\bar{\beta} < a$.[9]

One new feature is that if $\underline{s} > \hat{s}$ and $a - \bar{\beta}$ is large, then $\lambda^{*} = \lambda^{**}$ for all μ so that the network never induces paid content. This is developed in appendix.

[9]The proof is simpler because V^f decreases with λ.

6 Category pricing and screening

Until now we assumed that the network proposed a unique linear tariff. We now discuss the possibility to achieve second-degree discrimination by offering a menu of linear tariffs. Obviously, there is no possibility to discriminate between different CPs without inducing some differential consumptions. Indeed, if consumers are not affected by the choices of the CPs, the consumption q would be the same for all CPs and they would always opt for the smallest price s. However, the network may try to raise its profits and the value offered to consumers by combining a higher price for the CPs with a lower price for consumers. Content producers eager to generate traffic (due to low β as in our model or high benefits a) may then opt for this option. The advantage for the network may be not only to extract more rent from CPs but also to induce more efficient levels of consumption. We define this strategy as "category pricing".

Category pricing: The network proposes two tariffs (s_H, r_H) and (s_L, r_L). The CP chooses a tariffs, the consumer observes the tariff and consumes.

Category pricing then amounts to define several classes or categories, that we denote H and L. Content providers choose which category they want to belong to and this information is transmitted to the consumers. The tariffs then depend on the category.

To analyze this strategy in a concise manner we assume that the content can only be free and the benefits cover the traffic cost of the LL content:

Assumption 2 *μ is large and $\underline{\beta} < a$.*

Notice that if the network succeeds in inducing the LL and the HL content providers to choose different categories, then consumers should eventually realize that the average load is different for the two categories. They will thus adapt their behavior to the price of the category but also to the load in the category. This interaction between screening on one side and signalling on the other side is the difference between category pricing and a standard screening model.

We then define a revealing allocation for category pricing $\{(s_H, r_H), (s_L, r_L)\}$ as an allocation with the two properties below:

i) Consumers anticipate that the load is $\bar{\beta}$ for the category H and $\underline{\beta}$ for the category L;

ii) The HL (resp. LL) content providers choose category H (resp. L).

The first condition imposes that the consumers perfectly anticipate the traffic load by observing the category (hence a rational expectation equilibrium). This implies that we have consumptions $D\left(r_H\bar{\beta}\right)$ and $D\left(r_L\underline{\beta}\right)$ in the categories H and L respectively. Then we have a revealing tariff if the tariffs induce participation

$$a \geq s_H\bar{\beta} \text{ and } a \geq s_L\underline{\beta}$$

and the following incentive compatibility condition holds:[10]

$$\begin{aligned} \left(a - s_H\bar{\beta}\right) D\left(r_H\bar{\beta}\right) &\geq \left(a - s_L\bar{\beta}\right) D\left(r_L\underline{\beta}\right) \\ \left(a - s_L\underline{\beta}\right) D\left(r_L\underline{\beta}\right) &\geq \left(a - s_H\underline{\beta}\right) D\left(r_H\bar{\beta}\right) \end{aligned}$$

The value is then (recall that U is the representative consumer utility function):

$$\lambda\left[U\left(D\left(r_L\underline{\beta}\right)\right) + \left(s_L\underline{\beta} - \underline{\beta}\right) D\left(r_L\underline{\beta}\right)\right] + (1-\lambda)\left[U\left(D\left(r_H\bar{\beta}\right)\right) + \left(s_H\bar{\beta} - \bar{\beta}\right) D\left(r_H\bar{\beta}\right)\right]$$

When screening, the network maximizes the value under the participation constraints and the incentive compatibility constraints. To characterize the optimal tariff, we make the following change of variables:

$$q_H = D\left(r_H\bar{\beta}\right), \quad q_L = D\left(r_L\underline{\beta}\right), \quad S_H = s_H q_H, \quad S_L = s_L q_L$$

[10] Notice that a revealing allocation may induce $s_H = s_L$ as long as $r_H = r_L = 0$. In this case the CPs are indifferent between revealing their types to consumers or not. Thus they can choose category H or L depending on their type although there is no payoff difference.

Then the program of the network becomes

$$\begin{aligned} V &= \max \lambda \left[U(q_L) + \underline{\beta} S_L - \underline{\beta} q_L\right] + (1-\lambda)\left[U(q_H) + \bar{\beta} S_H - \bar{\beta} q_H\right] \\ st\ aq_H - S_H\bar{\beta} &\geq aq_L - S_L\bar{\beta} \\ aq_L - S_L\underline{\beta} &\geq aq_H - S_H\underline{\beta} \\ aq_H &\geq S_H\bar{\beta} \text{ and } aq_L \geq S_L\underline{\beta} \end{aligned}$$

From standard contract theory arguments, the solution involves $aq_H = S_H\bar{\beta}$ and $aq_L - S_L\underline{\beta} = aq_H - aq_H\underline{\beta}/\bar{\beta}$. Using these properties the program reduces to

$$V = \max_{q_L, q_H} \lambda \left[U(q_L) + aq_L - aq_H \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} - \underline{\beta} q_L\right] + (1-\lambda)\left[U(q_H) + aq_H - \bar{\beta} q_H\right]$$

where the quantities are restricted to the range $U'(q) \geq 0$. The solution is then

Proposition 5 *With free content and category pricing, the network charges tariffs:*

$$\begin{aligned} s_H &= \bar{s},\ r_H = \max\left(1 - \bar{s} + \frac{\lambda}{1-\lambda}\frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}}\bar{s}, 0\right) \\ s_L &= \underline{s}\left(1 - \frac{q_H}{D(0)}\frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}}\right),\ r_L = 0. \end{aligned}$$

Proof. The optimum is at

$$\begin{aligned} U'(q_L) &= 0 \\ U'(q_H) &= \max\left\{\bar{\beta} - a + \frac{\lambda}{1-\lambda} a \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}}, 0\right\} \end{aligned}$$

We have $q_H \geq q_L$,and all incentive constraints and participation constraints are satisfied. ■

The first question that arises is whether it is indeed optimal for the network to induce separation. The alternative is a pooling allocation at $s = \bar{s}$ and $r = \max(1 - \bar{s}, 0)$. In this pooling allocation, consumers do not make any inference and base their consumption on the average load β^e. Because less information is transmitted to the consumers, such a pooling involves an efficiency loss. We show below that in addition, separation helps the network extracting the rent of the LL content. Defining $\bar{\lambda}$ as the solution of

$$\bar{\beta} - a + \frac{\bar{\lambda}}{1-\bar{\lambda}} a \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} = \bar{p}$$

we obtain:

Lemma 7 *If $\lambda < \bar{\lambda}$, the network always prefers to use category pricing than a single tariff. If $\lambda \geq \bar{\lambda}$, the network excludes the HL content.*[11]

Proof. Suppose first that $\bar{s} \geq 1$. Then the price for consumers is $r = 0$ so that the allocation implemented with a single tariff is $q_H = q_L = D(0)$ and $s_H = s_L = \bar{s}$. This allocation is in the set of revealing allocations (see footnote 6), and is thus dominated by the optimal category pricing.

Suppose now that $\bar{s} < 1$. The pooling allocation is $s = \bar{s}$ and $r = 1 - \bar{s}$ with demand $D((1-\bar{s})\beta^e)$. Consider the following revealing allocation

$$\begin{aligned} s_H &= \bar{s};\ q_H = D\left((1-\bar{s})\bar{\beta}\right) \\ s_L &= \underline{s}\left(1 - \frac{q_H}{q_L}\frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}}\right);\ q_L = D\left(\max(1 - s_L, 0)\underline{\beta}\right) \end{aligned}$$

We first claim that such an allocation exists and $s_L > \bar{s}$. To see that define $\sigma(s)$ as

$$\sigma(s) = \underline{s}\left(1 - \frac{q_H}{D\left(\max(1-s,0)\underline{\beta}\right)}\frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}}\right)$$

[11] Notice that $\bar{\lambda} > \lambda^*$. Indeed the fact that the network chooses exclusion of HL when discrimination is allowed implies that it prefers that to a bunching at $\bar{s}$.

This mapping is non-decreasing with s, with value $\sigma(\bar{s}) > \bar{s}$ and maximal value $\sigma(1) = \sigma(\underline{s}) = \underline{s}\left(1 - \frac{q_H}{D(0)}\frac{\bar{\beta}-\underline{\beta}}{\underline{\beta}}\right) < \underline{s}$.Hence there exists $s_L > \bar{s}$ solution of $\sigma(s_L) = s_L$ which is the value in the revealing allocation.

The value V under pooling is (by convexity of S)

$$\begin{aligned} S\left((1-\bar{s})\beta^e\right) &< \lambda S\left((1-\bar{s})\bar{\beta}\right) + (1-\lambda) S\left((1-\bar{s})\underline{\beta}\right) \\ &< \lambda S\left((1-\bar{s})\bar{\beta}\right) + (1-\lambda)\left[S\left((1-\bar{s})\underline{\beta}\right) + \left(-\bar{s}\underline{\beta} + s_L\underline{\beta}\right) D\left(r\underline{\beta}\right)\right] \\ &< \lambda S\left((1-\bar{s})\bar{\beta}\right) + (1-\lambda)\max_r \left\{S\left(r\underline{\beta}\right) + \left(r + s_L\underline{\beta} - \underline{\beta}\right) D\left(r\underline{\beta}\right)\right\}. \end{aligned}$$

The latter value is the value under the revealing allocation.

Finally the category pricing induces a zero consumption of HL content for $\lambda \geq \bar{\lambda}$, which amounts to exclusion of this type of content. ■

Thus category pricing is clearly preferred by the network. From a welfare perspective, if we compare with the absence of second degree price discrimination we find that:

Corollary 5 *Comparing category pricing with the case of a uniform tariff*

- *if* $\lambda > \bar{\lambda}$: *then welfare is unchanged;*
- *if* $\bar{\lambda} > \lambda > \lambda^*$: *category pricing increases welfare, the rent of the LL content is larger and the HL content stays in the market.*
- *if* $\lambda < \lambda^*$ *and* $\bar{s} > 1$: *category pricing decreases welfare;*
- *if* $\lambda < \lambda^*$ *and* $\bar{s} < 1$: *category pricing increases welfare if* λ *is small enough*

Proof. See appendix ■

As it is now standard in the analysis of price-discrimination, category pricing raises welfare if it avoids the exclusion of the high load content. The new feature is that the LL content also benefits from the discrimination. The reason here is that only second-degree discrimination is allowed. Thus when allowing the HL content to stay with a low s (when $\bar{\lambda} > \lambda > \lambda^*$) the network needs to leave some rent to the LL content that were not needed with exclusionary uniform prices.

When there is no exclusion with uniform price, category pricing reduces welfare if it leads to higher price r for consumers which occurs when $\bar{s} > 1$. Indeed in this case the situation with uniform prices is efficient (with $r^* = 0$) while discrimination leads the network to reduce the volume for HH traffic by charging $r_H > 0$ for the motive of extracting more rent from the LL content.

When the consumers face a positive price with uniform tariffs, the effect of screening is more ambiguous as $r_L < r^* < r_H$. The price is efficient for the LL content but higher for the HH content. If there is little LL content, the distortion of r_H is small and the former effect dominates. But for larger values of λ we could not conclude.

7 Conclusion

This paper has investigated the impact of missing prices for the efficient pricing of network capacities. It has been shown that, when consumers control their consumption but are not aware of the induced effect, a direct or indirect signal should be send to them. In the standard setting with paid content, this signal is sent through the price chosen by the content producer and the network should post a positive price for content to transmit this information for consumers. When the content can choose to be either free or paid, the choice of price influences both the business model and the efficiency of exchange. Then cost sharing between the network and content producers has two benefits: it raises efficiency for the paid content and reduces the price charged to consumers due to a waterbed effect. These benefits need to be compared with potential costs in term of inefficient choice of business model or exclusion of some content. The analysis suggests that some partial cots sharing is efficient.

Even if we mainly focused on the case of a unique linear tariff, we extended our inquiry by allowing the network to propose a menu of tariff, among which each content provider must choose. By letting each content provider choosing not only its own price but also the price paid by their consumers, category pricing avoids the exclusion of the traffic intensive content and raises the volume for the less traffic intensive content. However it does not always increase welfare because the network may reduce excessively he volume of traffic induced in some category to raise the revenue in another category.

Two natural extensions can be envisioned. First, by allowing the use of hookup fees, the network objective was very close (and sometimes) similar to social welfare. It would be interesting to study the choice of tariff in a context of a monopoly with only a linear price for consumers. We can easily foresee that while the network would choose levels above marginal cost, some of our insights would still hold. In particular allowing the network to charge a positive price for content would result in lower prices for consumers. Second, we assumed that each content was competing with others so that the profit was reduced to zero in the paid-system. Allowing some market power on the content producers' side would change not only the choice of tariff by the network but also the way the information on the cost load is transmitted to the consumers. We hope to develop those issues in future research.

References

[1] Anderson S. and A. de Palma (forthcoming): "Competition for Attention in the Information (Overload) Age", *RAND Journal of Economics*

[2] Bourreau, Kourandi and Valletti (2012) "Net Neutrality with Competing Internet Platforms", mimeo telecom-ParisTech.

[3] Cañon C, 2009. " Regulation Effects on Investment Decision in Two-Sided Market Industries : The Net Neutrality Debate ", TSE mimeo.

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

[5] Economides and Hermalin (2012): "The Economics of Network Neutrality", New York University Law and Economics Working Papers

[6] Economides N. and J. Tåg, (2012).“Net Neutrality on the Internet : A two-sided Market Analysis ”, *Information Economics and Policy*, vol. 24 (2012) pp. 91–104

[7] Flakinger J. (2007), "Attention economies", Journal of Economic Theory, Volume 133, Issue 1, March 2007, Pages 266–294

[8] Falkinger J. (2008), "Limited Attention as a Scarce Resource in Information-Rich Economies", *The Economic Journal*, Volume 118, Issue 532, pages 1596–1620.

[9] Grafenhofer (2011): "Price Discrimination and the Hold-Up Problem: A Contribution to the Net-Neutrality Debate", Mimeo TSE

[10] Hermalin B. and M.L. Katz, 2007. “The Economics of product-line restrictions with an application to the network neutrality debate ”, *Information Economics and Policy*, Vol. 19, 215-248.

[11] Hermalin B. and M.L. Katz, 2009. “Information and the hold-up problem ”, *The Rand Journal of Economics*, Vol. 19, 215-248.

[12] Rochet J.C. and J.Tirole, 2003. “Platform Conpetition in Two-Sided Markets ”, *Journal of the European Economic Association*, Vol. 1, 990-1029.

[13] van Zandt T. (2004), "Information Overload in a Network of Targeted Communication", *RAND Journal of Economics*, 35(3), 542-560.

A Appendix A

Proof of proposition 2

In each range we have $r = \arg\max_r S(r\phi) + (r + s - 1)\,\phi D(r\phi)$, thus decreases with s. Moreover (with M denoting the volume of content)

$$\frac{\partial W}{\partial s} = M\left[(r\phi + a - \phi)\,D'(r\phi)\right]\phi\frac{\partial r}{\partial s}$$

which is positive because if $s < \bar{s}$

$$\begin{aligned} r\phi + a - \phi &= \max(1 - s, 0)\,\beta^e + a - \beta^e \\ &= \max(\bar{s} - s, \bar{s} - 1)\,\beta^e + \bar{s}\left(\bar{\beta} - \beta^e\right) > 0, \end{aligned}$$

while if $\bar{s} < s \leq \underline{s}$

$$\begin{aligned} r\phi + a - \phi &= \max(1 - s, 0)\,\underline{\beta} + a - \underline{\beta} \\ &= \max(\underline{s} - s, \underline{s} - 1)\,\underline{\beta} > 0. \end{aligned}$$

Hence the optimum is reached at the boundaries, $\bar{s}$ or $\underline{s}$.

Exclusion is not optimal if λ is small, while it is optimal if λ is close to 1 when $\bar{s}$ (because $\bar{R}$ if suboptimal when $\lambda \simeq 1$).

For $\lambda \leq \lambda^*$, (using $\bar{s}\beta^e < a$)

$$\begin{aligned} \underline{W} - \bar{W} &< \lambda\left[S(\underline{R}) + \left(\underline{R} + a - \underline{\beta}\right)D(\underline{R})\right] - \left[S\left(\bar{R}\right) + \left(\bar{R} + \bar{s}\beta^e - \beta^e\right)D\left(\bar{R}\right)\right] \\ &< 0 \end{aligned}$$

By continuity this is true for λ larger but close to λ^*. Hence $\bar{\lambda}^w > \lambda^*$.

For λ close to 1 and $\bar{s} < 1$:

$$\begin{aligned} \underline{W} - \bar{W} &\simeq \left[S(\underline{R}) + \left(\underline{R} + a - \underline{\beta}\right)D(\underline{R})\right] - \left[S\left(\bar{R}\right) + \left(\bar{R} + a - \underline{\beta}\right)D\left(\bar{R}\right)\right] \\ &> 0 \end{aligned}$$

because $\underline{R}$ maximizes social surplus. Hence $\underline{\lambda}^w < 1$.

For λ close to 1 and $\bar{s} \geq 1$:

$$\begin{aligned} \underline{W} - \bar{W} &= \lambda \left[S(0) + \left(a - \underline{\beta} \right) D(0) \right] - \left[S(0) + \left(a - \beta^e \right) D(0) \right] \\ &= -(1-\lambda) \left[S(0) + \left(a - \bar{\beta} \right) D(0) \right] < 0 \end{aligned}$$

Hence $\underline{\lambda}^w < 1$.

The slope of the surplus differential is

$$\frac{\partial \left(\underline{W} - \bar{W} \right)}{\partial \lambda} = \left[S\left(\underline{R} \right) + \left(\underline{R} + \left(\underline{s} - 1 \right) \underline{\beta} \right) D\left(\underline{R} \right) \right] - \bar{s} \left(\bar{\beta} - \beta^e \right) D'\left(\bar{R} \right) \frac{\partial \bar{R}}{\partial \lambda} + D\left(\bar{R} \right) \left(\underline{\beta} - \bar{\beta} \right)$$

This is linear if $\bar{s} > 1$ as $\frac{\partial \bar{R}}{\partial \lambda} = 0$. In this case $\underline{W} - \bar{W}$ is increasing and $\bar{\lambda}^w = \underline{\lambda}^w = \lambda^w$.

For $\bar{s} < 1$, we have $\bar{R} = \beta^e \max(1 - \bar{s}, 0)$:

$$\begin{aligned} \frac{\partial^2 \left(\underline{W} - \bar{W} \right)}{\partial \lambda^2} &= \bar{s} \left(\underline{\beta} - \bar{\beta} \right) D'\left(\bar{R} \right) \left(\underline{\beta} - \bar{\beta} \right) (1 - \bar{s}) - \bar{s} \left(\bar{\beta} - \beta^e \right) D''\left(\bar{R} \right) \left(\underline{\beta} - \bar{\beta} \right)^2 (1 - \bar{s})^2 \\ &\quad + D'\left(\bar{R} \right) \left(\underline{\beta} - \bar{\beta} \right)^2 (1 - \bar{s}) \\ &= \left[\bar{s} D'\left(\bar{R} \right) - \bar{s} \left(\bar{\beta} - \beta^e \right) D''\left(\bar{R} \right) (1 - \bar{s}) + D'\left(\bar{R} \right) \right] \left(\underline{\beta} - \bar{\beta} \right)^2 (1 - \bar{s}) \end{aligned}$$

which is negative if $D'' \geq 0$. In this case $\underline{W} - \bar{W}$ is concave which implies again $\bar{\lambda}^w = \underline{\lambda}^w = \lambda^w$. ■

Proof of lemma 3

For λ_0 we have

$$\begin{aligned} \frac{S\left(\beta_0^e \right) + aD\left(\beta_0^e \right)}{\lambda_0} &= \max_{R \geq 0} \left[S(R) + \left(R + a - \underline{\beta} \right) D(R) \right] \\ \textit{with } \beta_0^e &= \lambda_0 \underline{\beta} + (1 - \lambda_0) \bar{\beta} \end{aligned}$$

When a is large, thus reduces to

$$\frac{S\left(\beta_0^e \right) + aD\left(\beta_0^e \right)}{\lambda_0} = S(0) + \left(a - \underline{\beta} \right) D(0)$$

which leads to an approximate value when a goes to infinity, solution of

$$\frac{D\left(\beta_0^e\right)}{\lambda_0} \simeq D\left(0\right)$$

while $\lambda^* \simeq \bar{\lambda} = \frac{\bar{\beta}}{2\bar{\beta}-\underline{\beta}}$. Notice that $D\left(\beta^e\right) - \lambda D\left(0\right)$ is convex in λ, positive at $\lambda = 0$ and negative at $\lambda = 1$. Hence λ_0 is uniquely defined and $\lambda^* < \lambda_0$ if $D\left(\beta^{e*}\right) - \lambda^* D\left(0\right) > 0$. Hence $\lambda^* < \lambda_0$ for a large if

$$D\left(\frac{\bar{\beta}^2}{2\bar{\beta}-\underline{\beta}}\right) > \frac{\bar{\beta}}{2\bar{\beta}-\underline{\beta}} D\left(0\right)$$

Suppose now that a is small. As a first-order approximation we have when a is close to 0.

$$\begin{aligned} D\left(\underline{\beta}\right) da &= \frac{\underline{\beta}}{\bar{\beta}} D\left(\underline{\beta}\right) da - S\left(\underline{\beta}\right) d\lambda^* \\ \frac{d\lambda^*}{da} &= \left(\frac{\underline{\beta}}{\bar{\beta}} - 1\right) \frac{D\left(\underline{\beta}\right)}{S\left(\underline{\beta}\right)}. \end{aligned}$$

Moreover λ_0 tends also to 1when a goes to γ and

$$\begin{aligned} S\left(\beta_0^e\right) + aD\left(\beta_0^e\right) &= \lambda_0 \max_{R \geq 0} \left[S\left(R\right) + \left(R + a - \underline{\beta}\right) D\left(R\right)\right] \\ D\left(\underline{\beta}\right) da - D\left(\underline{\beta}\right)\left(\underline{\beta} - \bar{\beta}\right) d\lambda_0 &= D\left(\underline{\beta}\right) da + S\left(\underline{\beta}\right) d\lambda_0 \\ \frac{d\lambda_0}{da} &= 0 \text{ at } a = 1. \end{aligned}$$

Hence for a small, $\lambda^* < \lambda_0$. ■

Proof of lemma 6

We have

$$\frac{\partial V}{\partial r} = \lambda\left(s-1\right)\underline{\beta}^2 D'\left(r\underline{\beta}\right)$$
$$+\left(1-\lambda\right)\left(r+s-1\right)\bar{\beta}^2 D'\left(r\bar{\beta} + \left[s\bar{\beta} + \mu - a\right]_+\right)$$

In particular $\frac{\partial V}{\partial r} > 0$ if $r + s < 1$ and thus we have $r + s \geq 1$.

$$\frac{\partial V}{\partial s} = \lambda \underline{\beta} D\left(r\underline{\beta}\right) + (1-\lambda) I_{s\bar{\beta}+\mu-a<0} \bar{\beta} D\left(r\bar{\beta}\right)$$
$$+ (1-\lambda) I_{s\bar{\beta}+\mu-a>0} (r+s-1) \bar{\beta}^2 D'\left((r+s)\bar{\beta}+\mu-a\right)$$

In particular $\frac{\partial V}{\partial s} - \frac{\partial V}{\partial r} > 0$ as long as $r > 0$ which implies that $r = 0$ because increasing s and decreasing r at $r + s$ constant is profitable if $r > 0$.

We also have $\frac{\partial V}{\partial s} > 0$ if $r = 0$ and $s \leq 1$. Hence $s > 1$.

We thus solve (using $s\bar{\beta} + \mu - a > 0$)

$$\begin{aligned} V^p &= \max_{\bar{s} \leq s \leq \underline{s}} \lambda \left[S(0) + (s-1)\underline{\beta} D(0)\right] \\ &\quad + (1-\lambda)\left[S\left(s\bar{\beta}+\mu-a\right) + (s-1)\bar{\beta} D\left(s\bar{\beta}+\mu-a\right)\right] \end{aligned} \tag{6}$$

with slope

$$\begin{aligned} \frac{\partial V}{\partial s} &= \lambda \underline{\beta} D(0) + (1-\lambda)(s-1)\bar{\beta}^2 D'\left(s\bar{\beta}+\mu-a\right) \text{ if } s\bar{\beta}+\mu-a < \bar{p} \\ &= \lambda \underline{\beta} D(0) \text{ if } s\bar{\beta}+\mu-a > \bar{p} \end{aligned}$$

Notice that $\frac{\partial^2 V}{\partial \lambda \partial s} > 0$ and thus the optimal value of s^p is increasing (but may be discontinuous). For λ large the slope is positive on the interval $[1, \underline{s}]$ and thus $s^p = \underline{s}$. When λ goes to zero the slope becomes negative for all s between 1 and $\frac{\bar{p}+a-\mu}{\bar{\beta}}$ so that s^p goes to 1.

To conclude, we notice that for $s\bar{\beta} + \mu - a < \bar{p}$

$$\frac{\partial^2 V}{\partial s^2} = (1-\lambda)\bar{\beta}^2 \left\{D' + (s-1)\bar{\beta} D''\right\}$$

We have

$$D' + (s-1)\bar{\beta} D'' = D'(p) + \left(p - \left[\bar{\beta}+\mu-a\right]\right) D''(p) \textit{ with } p = s\bar{\beta}+\mu-a \in \left(\bar{\beta}+\mu-a, \bar{p}\right).$$

If $\left(p - \left[\bar{\beta}+\mu-a\right]\right) D'(p)$ decreases on this range, the objective is concave on this range. We may then distinguish two cases:

- If $\underline{s}\bar{\beta} + \mu - a < \bar{p}$ then the objective is concave and the optimal s is unique and continuous in λ. This occurs when $\frac{\bar{\beta}}{\underline{\beta}} < 1 + \frac{\bar{p}-\mu}{a}$.

- If $\underline{s}\bar{\beta} + \mu - a > \bar{p}$ then the objective is not globally concave but it is on the range $s\bar{\beta} + \mu - a < \bar{p}$. The network then chooses between $\underline{s}$ and s in the concave range. It chooses $s\bar{\beta} + \mu - a < \bar{p}$ for $\lambda < \lambda^x$ and in this case the solution is unique and continuous in λ, except at λ^x. ■

Proof of proposition 3

We know from the proof of proposition 2 that $\lambda \left[S(0) + (\underline{s} - 1)\underline{\beta}D(0)\right]$ crosses V^f only once and from below at λ^*. As $V^p \geq \lambda \left[S(0) + (\underline{s} - 1)\underline{\beta}D(0)\right]$, the network chooses $s \geq \bar{s}$ if $\lambda > \lambda^*$.

To be more precise, let the optimal value of s be s^p. Then

$$\frac{\partial V^p}{\partial \lambda} = S(0) + (s^p - 1)\underline{\beta}D(0) - \left[S\left(s^p\bar{\beta} + \mu - a\right) + \left(s^p\bar{\beta} - \bar{\beta}\right)D\left(s^p\bar{\beta} + \mu - a\right)\right] > 0$$

because $s^p > 1$ and

$$\begin{aligned} S\left(s^p\bar{\beta} + \mu - a\right) + \left(s^p\bar{\beta} - \bar{\beta}\right)D\left(s^p\bar{\beta} + \mu - a\right) &< \max_p S(p) + \left(p - \left[\bar{\beta} + \mu - a\right]\right)D(p) \\ &< S(0) \end{aligned}$$

because $\bar{\beta} + \mu - a > 0$.

At $\lambda = 0$, $V^f = S\left(\bar{\beta} - a\right) > \hat{V}^p = S\left(\bar{\beta} + \mu - a\right)$ so that again $\lambda^{**} > 0$.

$V^f = S\left((1 - \bar{s})\beta^e\right)$ is convex and we know that it crosses $\lambda \left[S(0) + (\underline{s} - 1)\underline{\beta}D(0)\right]$ once and from above. We wish to show that V^f crosses $\hat{V}^p$ only once and from above. Defining $\hat{s}^p$ as the argument in $\hat{V}^p$:

$$\frac{\partial \hat{V}^p}{\partial \lambda} = S(0) + (\hat{s}^p - 1)\underline{\beta}D(0) - \left[S\left(\hat{s}^p\bar{\beta} + \mu - a\right) + \left(\hat{s}^p\bar{\beta} - \bar{\beta}\right)D\left(s^p\bar{\beta} + \mu - a\right)\right]$$

$$\frac{\partial^2 \hat{V}^p}{\partial \lambda^2} = \frac{\partial \hat{s}^p}{\partial \lambda}\left(\underline{\beta}D(0) - \left(s^p\bar{\beta} - \bar{\beta}\right)\bar{\beta}D'\left(s^p\bar{\beta} + \mu - a\right)\right) > 0.$$

where the latter in equality follows form convexity of S. Thus $\hat{V}^p$ is convex which is enough to conclude. But notice that

$$\begin{aligned}\hat{V}^p &> \lambda S(0) + (1-\lambda) S\left(\bar{\beta} + \mu - a\right) \\ \frac{\partial \hat{V}^p}{\partial \lambda} &> S(0) - S\left(\bar{\beta} + \mu - a\right)\end{aligned}$$

because $S\left(\hat{s}^p\bar{\beta} + \mu - a\right) + \left(\hat{s}^p\bar{\beta} - \bar{\beta}\right) D\left(\hat{s}^p\bar{\beta} + \mu - a\right) < S\left(\bar{\beta} + \mu - a\right)$.

As $\lambda S(0) + (1-\lambda) S\left(\bar{\beta} + \mu - a\right)$ is linear, smaller than V^f at $\lambda = 0$ and larger at $\lambda = 1$, V^f crosses $\lambda S(0) + (1-\lambda) S\left(\bar{\beta} + \mu - a\right)$ only once and from above at some $\hat{\lambda}$. Thus V^f and $\hat{V}^p$ can only cross for $\lambda < \hat{\lambda}$. But when $\lambda < \hat{\lambda}$, by convexity of V^f the slope is smaller than at $\hat{\lambda}$ which implies

$$\frac{\partial V^f}{\partial \lambda} < S(0) - S\left(\bar{\beta} + \mu - a\right) < \frac{\partial \hat{V}^p}{\partial \lambda}.$$

Hence V^f crosses $\hat{V}^p$ only once and from above, as well as $\lambda\left[S(0) + (\underline{s} - 1)\underline{\beta} D(0)\right]$. Thus the content will be free for λ below λ^{**}. ■

Case where $\bar{\beta} < a$

We show that if $\bar{s}$ is large even at μ, we have $\lambda^{**} = \lambda^*$.

At $\lambda = \lambda^*$, $\hat{V}^p$ decreases in μ and taking the limit when μ small

$$\begin{aligned}\lim_{\mu\to 0} \hat{V}^p &= \max_{\bar{s}\leq s\leq \hat{s}} \lambda^* \left[S(0) + (s-1)\underline{\beta} D(0)\right] \\ &\quad + (1-\lambda^*)\left[S\left(s\bar{\beta} - a\right) + (s-1)\bar{\beta} D\left(s\bar{\beta} - a\right)\right]\end{aligned}$$

Notice that the value of the RHS at $s = \bar{s}$ is $S(0) + (\bar{s} - 1)\beta^{e*} D(0) = V^f$. Thus $\lim_{\mu\to 0} \hat{V}^p > V^f$ if the optimum in the RHS is at $s > \bar{s}$ which holds if

$$\frac{\lambda^*}{1-\lambda^*} > \left(a - \bar{\beta}\right) \frac{D'(0)}{D(0)}.$$

Thus if $a - \bar{\beta}$ is large, we have $\lambda^x > \lambda^*$ for μ small. But if $a - \bar{\beta}$ is small, we will have $\lim_{\mu\to 0} \hat{V}^p = V^f$ which implies that $\hat{V}^p < V^f$ for all $\mu > 0$. In this case $\lambda^x < \lambda^*$ implying that $\lambda^{**} = \lambda^*$. ■

Proof of corollary 5

We need to compare

$$\begin{aligned} W^D &= \lambda\left[S(0)+\left(a-\underline{\beta}\right)D(0)\right]+(1-\lambda)\left[S(R_H)+\left(R_H+a-\bar{\beta}\right)D(R_H)\right] \\ \text{where } R_H &= \max\left(1-\bar{s}+\frac{\lambda}{1-\lambda}\frac{\bar{\beta}-\underline{\beta}}{\bar{\beta}}\bar{s},0\right)\bar{\beta} \end{aligned}$$

and

$$\begin{aligned} \bar{W} &= S\left(\bar{R}\right)+\left(\bar{R}+a-\beta^e\right)D\left(\bar{R}\right) \\ \bar{R} &= \max(1-\bar{s},0)\,\beta^e \end{aligned}$$

But

$$(1-\lambda)R_H=\max\left(1-\bar{s}+\frac{\lambda}{1-\lambda}\frac{\bar{\beta}-\underline{\beta}}{\bar{\beta}}\bar{s},0\right)(1-\lambda)\bar{\beta}.$$

If $\bar{s}>1$, then $\bar{W}=S(0)+(a-\beta^e)D(0)$ which is the second-best welfare (constrained by the non-negativity constraint on prices) and category pricing can only reduce total welfare.

If $\bar{s}<1$,

$$\begin{aligned} \bar{W} &= S\left(\bar{R}\right)+\left(\bar{R}+a-\beta^e\right)D\left(\bar{R}\right)=S\left(\bar{R}\right)+\lambda\frac{\bar{\beta}-\underline{\beta}}{\bar{\beta}}aD\left(\bar{R}\right) \\ \bar{R} &= \lambda\left(\underline{\beta}-a\frac{\underline{\beta}}{\bar{\beta}}\right)+(1-\lambda)\left(\bar{\beta}-a\right) \end{aligned}$$

compared with

$$\begin{aligned} W^D &= \lambda\left[S(0)+\left(a-\underline{\beta}\right)D(0)\right]+(1-\lambda)S(R_H)+\lambda\frac{\bar{\beta}-\underline{\beta}}{\bar{\beta}}aD(R_H) \\ \text{where } R_H &= \bar{\beta}-a+\frac{\lambda}{1-\lambda}\frac{\bar{\beta}-\underline{\beta}}{\bar{\beta}}a. \end{aligned}$$

$$\frac{\partial \bar{W}}{\partial \lambda} = \left(\bar{\beta} - \underline{\beta}\right) D\left(\bar{R}\right) - \lambda \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a D'\left(\bar{R}\right) \left(\bar{\beta} - \underline{\beta} - \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a\right).$$

$$\frac{\partial W^D}{\partial \lambda} = S(0) - S(R_H) + \left(a - \underline{\beta}\right) D(0) - \frac{\lambda}{1-\lambda} \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a D(R_H) + \frac{\lambda}{(1-\lambda)^2} \left(\frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a\right)^2 D'(R_H).$$

At $\lambda = 0$

$$\begin{aligned} \frac{\partial W^D}{\partial \lambda} &= S(0) - S\left(\bar{\beta} - a\right) + \left(a - \underline{\beta}\right) D(0) \\ &> \left(\bar{\beta} - a\right) D\left(\bar{\beta} - a\right) + \left(a - \underline{\beta}\right) D(0) > \left(\bar{\beta} - \underline{\beta}\right) D\left(\bar{\beta} - a\right) = \frac{\partial \bar{W}}{\partial \lambda} \end{aligned}$$

For the record notice that looking at second derivatives

$$\begin{aligned} \frac{\partial^2 \bar{W}}{\partial \lambda^2} &= -\left(\bar{\beta} - \underline{\beta}\right) D'\left(\bar{R}\right) \left(\bar{\beta} - \underline{\beta} - \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a\right) - \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a D'\left(\bar{R}\right) \left(\bar{\beta} - \underline{\beta} - \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a\right) \\ &\quad + \lambda \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a D''\left(\bar{R}\right) \left(\bar{\beta} - \underline{\beta} - \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a\right)^2 \end{aligned}$$

is positive if D is convex or linear.

$$\frac{\partial R_H}{\partial \lambda} = \frac{1}{(1-\lambda)^2} \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a$$

$$\begin{aligned} \frac{\partial^2 W^D}{\partial \lambda^2} &= D(R_H) \frac{1}{(1-\lambda)^2} \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a - \frac{1}{(1-\lambda)^2} \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a D(R_H) - \frac{\lambda}{(1-\lambda)^3} \left(\frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a\right)^2 D'(R_H) \\ &\quad + \frac{1+\lambda}{(1-\lambda)^3} \left(\frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a\right)^2 D'(R_H) + \frac{\lambda}{(1-\lambda)^2} \left(\frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a\right)^2 D''(R_H) \frac{1}{(1-\lambda)^2} \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a \end{aligned}$$

$$\frac{\partial^2 W^D}{\partial \lambda^2} = \frac{1}{(1-\lambda)^3} \left(\frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a\right)^2 \left(D'(R_H) + \frac{\lambda}{1-\lambda} \frac{\bar{\beta} - \underline{\beta}}{\bar{\beta}} a D''(R_H)\right)$$

So W^D is concave if the $\left(R + \bar{\beta} - a\right) D'(R)$ is non-increasing. In this case there exists a threshold $\underline{\lambda}$ below which PD increases welfare.

B Appendix B : Results of the GENERAL CASE MODEL

We now derive results for the case where (β, a, μ) are distributed on $\mathbb{R}_+^3$ with cdf F. Interpreting a as ads receipt net of distribution costs, we allow a to be negative to capture the existence of content that should be paid even at $s = 0$. The objective is to understand how the optimal price s is determined. For this we consider the situation where the regulator chooses s in a first step and then the network chooses r (which is not regulated).

We have a paid content with $p = s\beta + \mu - a$ if $s\beta > a$ and a free content otherwise. When p is above $\bar{p}$, the content is excluded. As before define $\phi(s) = E(\beta \mid s\beta \leq a)$ and $\chi(s, p) = E(\beta \mid p, s)$ for $p > \mu$ then the value V is:

$$\begin{aligned} V &= \Pr(s\beta \leq a)\left[S(r\phi(s)) + (r + s - 1)\phi(s)D(r\phi(s))\right] \\ &\quad + \Pr(s\beta > a)E\left\{S(r\chi(s, p) + p) + (r + s - 1)\beta D(r\chi(s, p)) \mid s\beta > a\right\} \end{aligned}$$

The first immediate result is:

Lemma 8 *For any regulated price s, the consumer congestion price is $r = \max(1 - s, 0)$*

Proof. We have

$$\begin{aligned} V &= \Pr(s\beta \leq a)\left[S(r\phi(s)) + (r + s - 1)\phi(s)D(r\phi(s))\right] \\ &\quad + \Pr(s\beta > a)E\left\{S(r\chi(s, p) + p) + (r + s - 1)\chi(s, p)D(r\chi(s, p)) \mid s\beta > a\right\} \end{aligned}$$

With this formulation we see that $r = \max(1 - s, 0)$ maximizes the term in bracket and each term in the expectation for paid content (because p is independent of r) ■

Then welfare is

$$\begin{aligned}W &= \Pr(s\beta \le a) E\{S(r\phi(s)) + (r\beta + a - \beta) D(r\phi(s)) \mid s\beta \le a\} \\ &\quad + \Pr(s\beta > a) E\{S(r\chi(s,p) + p) + (r + s - 1)\beta D(r\chi(s,p) + p) \mid s\beta > a\}\end{aligned}$$

Proposition 6 *The optimal regulated price s^R is less or equal than* 1.

Proof. Suppose $s \ge 1$, then

$$\begin{aligned}W &= \Pr(s\beta \le a) E\{S(0) + (a - \beta) D(0) \mid s\beta \le a\} \\ &\quad + \Pr(s\beta > a) E\{S(s\beta + \mu - a) + (s - 1)\beta D(s\beta + \mu - a) \mid s\beta > a\}\end{aligned}$$

If $s > 1$, reducing s to 1 has two effects. First for content such that $\beta > a$, the welfare increases as the content stays paid and the price $p = \beta + \mu - a$ is efficient for paid content. Second the content such that $s\beta > a > \beta$ becomes free. For this content the welfare changes from the paid welfare to the maximal welfare $S(0) + (a - \beta) D(0)$. Thus welfare increases. ■

Let us now consider $s < 1$. In this case we have

$$\begin{aligned}W &= \Pr(s\beta \le a) E\{S((1-s)\phi(s)) + (a - s\beta) D((1-s)\phi(s)) \mid s\beta \le a\} \\ &\quad + \Pr(s\beta > a) E\{S((1-s)\chi(s,p) + p) \mid s\beta > a\}\end{aligned}$$

Suppose that a has continuous distribution on $(\underline{a}, \bar{a})$ where $\underline{a} \le 0 < \bar{\beta} < \bar{a}$, assume also that a, β and μ are not correlated.

Proposition 7 *Starting from $s = 0$, an increase of s raises welfare for the contents that stay free but reduces it for the contents that change from free to pay. The total effect is positive if $F(0)$ and $f(0)/(1 - F(0))$ are small enough.*

Proof. We have

$$W = E\left\{\begin{array}{c}(1 - F(s\beta)) E\{S((1-s)\phi(s)) + (a - s\beta) D((1-s)\phi(s)) \mid s\beta \le a, \beta\} \\ + F(s\beta) E\{S((1-s)\chi(s,p) + p) \mid s\beta > a, \beta\}\end{array}\right\}$$

The slope is

$$\frac{\partial W}{\partial s} = E\left\{\begin{array}{c} \left(1-F\left(s\beta\right)\right)E\left\{-\frac{\partial\phi}{\partial s}D + aD'\left(\frac{\partial\phi}{\partial s}-\phi\right) \mid s\beta \le a,\beta\right\} \\ +F\left(s\beta\right)E\left\{-\left(-\chi+\beta+\frac{\partial\chi}{\partial s}\right)D \mid s\beta > a,\beta\right\} \\ +\beta f\left(s\beta\right)E\left\{S\left(\left(1-s\right)\chi+p\right)-\left[S\left(\left(1-s\right)\phi\right)+\left(a-s\beta\right)D\left(\left(1-s\right)\phi\right)\right] \mid s\beta = a,\beta\right\} \end{array}\right\}$$

At $s=0$, the price p is independent of β as well as $s\beta$, so that we have

$$\left.\frac{\partial W}{\partial s}\right|_0 = \left\{\begin{array}{c} \left(1-F\left(0\right)\right)E\left\{-\frac{\partial\phi}{\partial s}D + aD'\left(\frac{\partial\phi}{\partial s}-\phi\right) \mid 0 \le a\right\} \\ +F\left(0\right)E\left\{-\left(-\chi+\beta^e+\frac{\partial\chi}{\partial s}\right)D \mid 0 \ge a\right\} \\ +\beta^e f\left(0\right)E\left\{S\left(\chi+\mu-a\right)-S\left(\phi\right) \mid a=0\right\} \end{array}\right\}$$

Using the fact that at $s=0$ we have $\chi\left(0,p\right)=\phi\left(0\right)=\beta^e$

$$\begin{aligned} \left.\frac{\partial W}{\partial s}\right|_0 &= \left(1-F\left(0\right)\right)E\left\{-\frac{\partial\phi}{\partial s}D\left(\beta^e\right)+aD'\left(\beta^e\right)\left(\frac{\partial\phi}{\partial s}-\beta^e\right) \mid 0 \le a\right\} \\ &\quad +F\left(0\right)E\left\{-\frac{\partial\chi\left(0,\mu-a\right)}{\partial s}D\left(\beta^e+\mu-a\right) \mid 0 \ge a,\beta\right\} + \beta^e f\left(0\right)E\left\{S\left(\beta^e+\mu\right)-S\left(\beta^e\right)\right\} \end{aligned}$$

The first term is positive and the last term is negative. ■