Net Neutrality and Universal Service Obligations

Axel Gautier[*], Jean-Christophe Poudou[†] and Michel Roland[‡]

June 23, 2022

Abstract

This paper analyzes whether repealing net neutrality (NN) improves or decreases the capacity of a regulator to make internet service providers (ISPs) extend broadband coverage through universal service obligations (USOs). We model a two-sided market where a monopolistic ISP links content providers (CPs) to end users with a broadband network of a given bandwidth. A regulator determines whether to submit the ISP to NN or to allow it to supply paid priority (P) services to CPs. She can also impose a broadband USO to the ISP, i.e. she can mandate the broadband market coverage. We show that the greater is the network bandwidth, the more likely the repeal of net neutrality increases ISP profits and social welfare. Regulation can still be necessary, however, as there are bandwidth ranges for which the ISP would benefit from a repeal of NN while such a repeal is detrimental to society.

Keywords: Internet, Net Neutrality, Universal Service Obligations, Prioritization, Regulation

JEL: D21, K23, L12, L51, L96

[*] LCII, HEC Liege, University Liege, Belgium. Email: agautier@uliege.be

[†] MRE, MUSE, University of Montpellier, Montpellier, France. Email: jean-christophe.poudou@umontpellier.fr

[‡] CREATE, Département d'économique, Université Laval, Québec, Canada. Email: Michel.Roland@ecn.ulaval.ca

1 Introduction

Most countries impose two types of regulation on Internet Service Providers (ISPs): Net Neutrality (NN) and Universal Service Obligations (USOs). The first prohibits ISPs from "speeding up, slowing down or blocking Internet traffic based on its source, ownership or destination" (Kramer et al. [21]). It aims to promote investment, innovation and competition among content providers (CPs) and more generally, to ensure free speech (Katz [20]). The latter forces ISPs to cover a given percentage of the territory with a minimum broadband standard, set in terms of download and upload speeds. Its goal is to avoid a "digital divide" among citizens of different regions (McMenemy [23]).

Because of the growth of data intensive content on internet, ISPs argue that it is nowadays counterproductive to treat in the same way CPs that require high speed of transmission and that do not tolerate delays, like streaming, from those that are far less demanding on those counts, like emails. Peitz and Schuett [25] show that when contents have different sensitivities to delay, letting the ISP organize a paid prioritization service could improve welfare. As a result, there are debates within regulatory agencies and among academics on the ongoing relevance of NN, while the Government of the US has already repealed the NN rules in 2018. In contrast, there is a clear tendency to strengthen USOs almost everywhere in the world, including in the US (Garci-Calvo [18]).

A striking feature of the debates and economic analysis on Internet regulation is that they treat NN and USOs policies independently. On the one hand, the growing literature on NN studies its impact on social welfare, content innovation and network investment (Calzada and Tselekounis [7]). The literature gives two interpretations to NN: in the first, the ISPs cannot charge CPs, i.e. NN is interpreted as a zero price rule on the CP side; in the second, ISPs cannot offer quality differentiated access to CPs, implying that they cannot prioritize some CPs' traffic, and consequently they cannot ask for a payment to prioritized content. Although economic models differ with respect of the assumed market structure (monopoly, oligopoly, vertically integrated firms, etc.) for ISPs as well as for CPs, they are similar in their objectives to compare market equilibria, where each agent maximizes profit or utility, with and without NN. In other words, the only regulation considered is NN and the benchmark case is free market. This ignores the fact that, in reality, there generally exist USOs and subsidies for network extension in remote areas. Subsidizing mechanisms potentially change both the market size and the market configuration.

On the other hand, the early literature on USOs takes their existence as given and evaluate compensations schemes for the ISPs in their capacities to cover the cost burden imposed on the universal service provider (USP) without altering firms' competitive behavior in markets (Gautier and Wauthy [17], for instance). Models also analyze the impact on social welfare and on entry incentives of a uniform price constraint, which is often part of USOs and which prohibits the ISPs to offer tariffs that differ among markets, i.e. to price discriminate (Valletti et al [32], for instance).

The literature on USOs is mute on the NN both because discussions to repeal the latter are rather recent and that, under NN, one can analyze Internet USOs with the same methodological approach than the one applied to other industries already studied, such as electricity, natural gas or traditional telecommunications. Indeed, since CPs do not pay specific charges to ISPs, their uploads are virtually free and, consequently, investment incentives in networks for ISPs come only from the end users willingness to pay for the service, as is the case for the other industries. In other words, with the zero price rule, the ISPs look like one sided firms selling access to internet to consumers. However, this vision is misleading and the behavior of the CP side of the market must be taken into account, especially when the NN rules are repealed.

The separate treatment of NN and USOs is clearly unsatisfactory as both regulations impact directly on Internet service pricing and on the incentives to invest in broadband networks. To evaluate fully the global impact of the repealing of NN on the industry performance in terms of pricing and investment, one has to consider whether it increases or decreases the regulator's capacity to extend USOs. This amounts to the question of whether a regulator who wishes to extend the network above the firms' profit maximizing coverage is able or not to capture to that end a greater slice of the industry rent following the repealing of NN.

In this paper, we analyze both NN and USO in a single model. We consider a two-sided market where a monopolistic ISP can install a broadband network of a given bandwidth at different locations of a country. A regulator can determine whether to submit the ISP to NN or to allow it to supply paid priority (P) services to CPs with the objective to maximize market coverage or, provided it has enough instruments, to maximize welfare. Prioritization gives the opportunity to the ISP to obtain revenue from the prioritized content providers. However, it tilts consumption towards these prioritized CPs and, as a result, can affect adversely consumers if they have strong preferences for the non-prioritized service. The overall impact depends on the network bandwidth. The greater is the bandwidth, i.e. the greater is the network data transfer rate, the lesser is the impediment of priority on non-prioritized content and the more likely it is that prioritization will increase total consumption, ISP profit and/or social welfare. However, because the detrimental effect of prioritization on non-prioritized content impacts on welfare but not on the ISP profit, there is a bandwidth range over which the repeal of NN increases the ISP profit while it decreases welfare. A regulation mandating net neutrality is then called for.

The intuitive property that the P regime favors the prioritized content is a basic feature of the NN literature. Across models, however, this property is ensured by using different mechanisms to implement prioritization. On the one hand, in most models, prioritization consists of distinguishing content types by different waiting times in a standard M/M/1 queue system.[1] For tractability, the consumers' utility function is separable in each content consumption and in average download time. The marginal utility of content consumption is constant and consumers do not face any constraint, so that they absorb any increase in content that is supplied. We characterize these models as (content) supply driven. In such models, download times are endogenous and an increase in bandwidth capacity tends to decrease the waiting times gap between contents under priority "because the marginal reduction in waiting time for the fast lane from capacity expansion decreases as the capacity level becomes high".[2] An implicit assumption behind this result is that bandwidth does not impact directly on aggregate content demand.

On the other hand, Economides and Hermalin [15] model priority as a division of the bandwidth into sub-bandwidths with different capacities. This defines the time necessary to download each content. The utility function is quasi-linear between a numeraire and internet content consumption and marginal utility of each content depends on download time. In contrast with supply-driven models, aggregate content demand depends directly on bandwidth capacity, so we characterize this model as demand driven. Because of quasi-linearity and the fact that the price charged for consuming content is independent of the content provider, they obtain the following result: "[G]iven two alternative divisions of the total bandwidth, one is welfare superior to the other if and only

[1] See, for instance, Reggiani and Valletti [28], Choi and Kim, [11] and Choi et al. [10].

[2] Choi and Kim [12].

if it results in more content being carried in equilibrium than the other." (p. 609). An increase in bandwidth then immediately brings an increase of welfare as it necessarily allows an increased traffic. In other words, in Economides and Hermalin [15], the impact of an increase in bandwidth can be assimilated to an income effect in the standard demand theory, while this impact can be assimilated to a substitution in supply driven models.

In order to take into account both these substitution and income effects simultaneously, we introduce a consumer time budget constraint where latency, the delay cost of downloading in terms of time, plays the role of price, and available bandwidth, the role of income. As a result, end users utility is only indirectly affected by prioritization, as is the case in Reggiani and Valletti [28]. However, this indirect impact goes through the demand side of the market rather than through the supply side, so that an increase in bandwidth brings an "income effect", i.e. the idea the consumer can use more data in general at a given download speed. We also assume a Cobb-Douglas utility function so that marginal utility of a particular content is not independent of the consumption of another content, as in the quasi-linear case. In contrast to Economides and Hermalin [15], this allows for cases where social welfare decreases even though more content is consumed in total and, for our purpose, the possibility that the expansion of network coverage comes with a decrease of the utility of existing consumers, as it is often the case with USO.

With respect to the USO literature, we use the standard framework in which a regulator determines the extent of a total market that the ISP network must cover, acknowledging the fact that some of the sub-markets are not profitable because consumers, in spite of having the same preferences over the network services, are heterogeneous with respect to their connection costs to the network. Generally in this large literature,[3] extension of service beyond the profit maximizing coverage must be financed through the industry profit. A recurrent theme is to evaluate the welfare impact of a uniform pricing constraint, which is a ban on third-degree price discrimination. To our knowledge, only one-sided markets have been analyzed. We rather consider a pricing constraint in a two-sided market, net neutrality, which is a ban on third-degree price discrimination on one side of the market, the content providers.

Investment in network capacity in our model corresponds to the extension of market coverage as in the USO literature and not to the increase in bandwidth as in the NN literature. We thus refer

[3]Early contributions are Anton et al. [1] and Valletti et al. [32].

to a case where bandwidth is primarily determined by the current state of technology, and accordingly, we assume it is exogenous. Historically, the interaction of technological improvement (from copper networks to fiber, for instance) and content data transmission requirements (from emails to streaming, for instance) has resulted in a ever increasing minimum standard for bandwidth to be considered as part of a high-speed broadband service. For instance, the FCC broadband definition has evolved from 200/200 Kbps download/upload speeds, to 4/1 Mbps in 2010 and then to 25 Mbps/3 Mbps in 2015,[4] and this is probably called for a revision soon.[5] Broadband definitions also vary across countries.[6] At the same time, most countries share the "Biden Administration's commitment to deploying affordable, high-speed broadband across the country to help bridge America's digital divide and remedy persistent digital inequities" (Bennett et al [3]). We show that the regulatory framework that is the most efficient to reach the common goal of a universal broadband coverage depends crucially on the network bandwidth that is envisioned. This fact could be overlooked as long as internet traffic was fairly homogeneous in terms of bandwidth requirements, so that the NN debate could be made independently of the establishment of USO. But its importance should increase as the consumption patterns vary in time and across countries.

In the next section, we present the model of the two-sided internet market that we analyze and we specify the way net neutrality and prioritization are defined and implemented. We also derive end-user demands of CP contents under both net neutrality and prioritization. In section 3, we perform the comparative statics between net neutrality and prioritization for a given market coverage and we present the benchmark cases of welfare and profit-maximizing coverages. Section 4 provides the core results on the choice between net neutrality and prioritization as well as on the determination of market coverage in function of bandwidth. For ease of presentation, these results

[4]See BroadbandNow [6].

[5]A number of signals go in this direction. For instance, as soon as in 2016, FCC Commissioner Jessica Rosenworcel [29] claims: "I am proud I was the first to call for a new broadband standard of 100 Megabits. I think anything short of that shortchanges our children, our digital economy, and our future". On its website, Verizon [33]: "If you love to stream HD videos, download large files and enjoy multiplayer gaming, you may want to consider speed plans of 100 Mbps and above". In 2021, a bipartisan group of four senators wrote an open letter urging to "update federal broadband program speed requirements to reflect current and anticipated 21st century uses" (Bennet et al [3]).

[6]For instance, the EU defines a 30 Mbps download speed as fast broadband and a 100 Mbps as ultrafast broadband (Bourreau [5]). Coverage targets are given in both terms. Canada sets broadband coverage targets in terms of 50/10 Mbps download/upload speeds (CRTC [8]).

are obtained under simplifying assumptions, but we present extensions and provide robustness checks in Sections 5 and 6, respectively. Although the extensions cover issues of great practical relevance for implementation of universal service obligations with or without net neutrality, they do not modify qualitatively our core results. The conclusion sums up the main results of our model.

2 Basic Model

We consider a two-sided market where a monopolistic ISP connects consumers to content providers in a country composed of a continuum of locations $n \in [0, \infty[$ that are ranked in increasing order of network deployment cost. A regulator oversees the ISP with the aim of maximizing social welfare given the regulatory tools it has in hands. We consider "regulatory frameworks" that differ by the use of either one or both of two different regulatory tools: (i) the enforcement of a "traffic management practice", which is a choice between net neutrality (N) and prioritization (P) and/or (ii) the imposition of universal service obligations, which is the choice of the ISP market coverage.

In this section, we describe a basic model that is sufficient to highlight the main trade-offs involved in the regulator's choices. Some simplifying assumptions are made in favor of readability and tractability. In section 5, we extend the model to take into account two factors of practical relevance, the participation constraint of the ISP and the possibility that CPs are foreign-owned. We relax more technical assumptions in section 6. In both cases, fundamental results of the basic model follow through.

2.1 Content Providers (CPs)

There are two types of content providers denoted by $j = 0, 1$. CPs value traffic on their websites or applications and they have an ad-sponsored business model. Each CP's total revenue is equal to the click probability times the revenue per click and we denote by a this expected benefit per unit of traffic. Operating costs are normalized to zero. Denoting by X_j the total traffic per location (in MB) of CP j, a CP's profit is $\Pi_j = anX_j$. We let $X \equiv X_0 + X_1$ be the total traffic per location.

Although consumers distinguish content from each CP through their preferences, in this basic model, CPs are homogeneous in terms of technology. However, in order to consider the impact of prioritization on content diversity, we introduce CP heterogeneity in section 6.3.

2.2 Internet Service Provider (ISP)

The ISP operates a network of bandwidth μ to link CPs to consumers in n locations. The maximum traffic that the ISP network can handle depends on bandwidth and the traffic management practiced in case of congestion. We use the standard M/M/1 queue system to model congestion, as it "is well known to be a very good approximation for the arrival process in real systems".[7] However, instead of using the M/M/1 model to determine, for a given bandwidth, the average content delivery delay in function of traffic, we determine traffic in function of delay. Waiting times, or its reciprocal, transmission speed, is then the quality of service advertised by the ISP.

Two traffic management practices are possible: net neutrality and prioritization.

Net Neutrality (N). With net neutrality, traffic is managed under a best effort service so that the ISP announces an average transmission speed in MB/s. This average speed turns out to be the reciprocal of the average waiting time $\bar{\omega}$, which is given by the M/M/1 queue model:

$$\bar{\omega} = \frac{1}{\mu - \lambda X}$$

where λ is the frequency (in s^{-1}) of data transmission to the network, which we assume identical across contents. Data arriving at a speed exceeding λX would involve an infinite waiting time and is therefore considered as not being served by the network. The ISP announces the average speed of transmission $\frac{1}{\bar{\omega}}$ as its quality of service under net neutrality. For bandwidth μ and the normalized quality of service $\bar{\omega}$ normalized to 1, the network is then able to support total data transmission $X = \frac{\mu - 1}{\lambda}$. We also normalize λ to 1, so that $\mu - 1$ represents the data transmission capacity of the system.[8] The network has thus a capacity constraint given by:

$$X_0 + X_1 = \mu - 1 \tag{1}$$

Prioritization (P). With a given bandwidth μ, the ISP can alternatively route traffic with a prioritization system under which waiting times are determined as if half the traffic $\mu - 1$ observed under neutrality is given precedence in case of congestion.[9] In other words, waiting times are

[7] Choi and Kim [11], p. 452. The M/M/1 queue system is also used in Choi et al. [10], Reggiani and Valletti [28], Choi and Kim [12], Bourreau et al. [4], and Kramer and Weiwiorra [22].

[8] Depending on the context, variable μ can then be referred either to bandwidth (in MB/s) or to data transmission capacity (in MB).

[9] Once waiting times are determined, the actual traffic that is prioritized will be determined endogenously.

defined with an equal endowment of capacity that is a priori allocated to the priority traffic and the non-priority class.[10]

Instead of posting an average speed $\frac{1}{\bar{\omega}} = 1$, the ISP announces priority speed $\frac{1}{\omega_0} > 1$ and "regular" speed $\frac{1}{\omega_1} < 1$ that result from the M/M/1 queue:[11]

$$\begin{aligned} \omega_0 &= \frac{1}{\mu - \frac{1}{2}(\mu - 1)} = \frac{1}{\frac{1}{2}(\mu + 1)} \\ \omega_1 &= \mu\omega_0 \end{aligned} \tag{2}$$

These speeds must still meet the overall average delay $\bar{\omega} = 1$: as a result, consumption levels X_0 and X_1 under prioritization must be such that waiting times ω_0 and ω_1 weighted by the consumption shares in transmission equal 1:[12]

$$\omega_0 \cdot \frac{X_0}{\mu - 1} + \omega_1 \cdot \frac{X_1}{\mu - 1} = \bar{\omega} = 1 \tag{3}$$

Using (2) and multiplying both sides by $\frac{\mu-1}{\omega_0}$, this can be written as:

$$X_0 + \mu X_1 = \frac{1}{2}\left(\mu^2 - 1\right) \tag{4}$$

Note that, by construction, consumption vector $(X_0, X_1) = \left(\frac{1}{2}(\mu - 1), \frac{1}{2}(\mu - 1)\right)$ is feasible under both traffic management practices.

Cost and Revenue

The cost of establishing a network that covers markets $[0, n]$ with a bandwidth μ is:

$$C(n, \mu) = \frac{1}{2} c\mu n^2 \tag{5}$$

so that the marginal cost of coverage is increasing with bandwidth. Note that because a bandwidth μ has a maximum data transmission capacity of $\mu - 1$, this function is defined on $\mathbb{R}_+ \times [1, \infty)$. $C(n, 1)$ can thus be interpreted as a fixed cost of serving n markets, as the network must install bandwidth $\mu = 1$ before being able to transmit any data in a finite time.

[10] This initial endowment is arbitrary and made for readability. In section 6.1, we show that qualitative results are unchanged if we use any share ρ that does not exceed the consumption share observed under neutrality of the content to be prioritized without neutrality.

[11] In view of the fact that CPs are homogeneous in this basic model, the choice of content to be prioritized is arbitrary at this stage. We provide a rationale for the choice of content to be prioritized in Section 6.3.

[12] An equivalent interpretation is to say that X_0 and X_1 must meet capacity constraint $\omega_0 X_0 + \omega_1 X_1 = \mu - 1$.

To cover these costs, the ISP charges a fixed fee to end users. The ISP can then extract completely the consumers' surplus. A fixed fee is also charged to providers of prioritized content under P traffic management, so that the ISP can appropriate the incremental profit brought to beneficiaries of prioritization. This assumes that prioritized CPs have no bargaining power. We relax this assumption in Section 6.2.

2.3 Consumers

There is a mass 1 of identical consumers in each location and we use the Cobb-Douglas function to represent their preferences:

$$U\left(X_0, X_1\right) = X_0^{\alpha} X_1^{\beta} \tag{6}$$

where $\alpha + \beta < 1$.

Total consumption is constrained by the data transmission capacity $\mu - 1$ and by the delay of transmission for each content ω_0^i and ω_1^i that prevails under traffic management $i = N, P$:

$$\omega_0^i X_0 + \omega_1^i X_1 = \mu - 1 \tag{7}$$

This can be interpreted as a standard budget constraint where delays play the role of prices and effective capacity, the role of income. Dividing both sides of (7) by ω_0^i, i.e. by considering the prioritized content as the numeraire, one obtains constraints (1) and (4) for the net neutrality and priority management techniques, respectively. Compared to (1), constraint (4) displays a higher "relative price" for non-prioritized content. Prioritization thus gives incentives to decrease the share of non-prioritized content to prioritized content in total consumption. In fact, the budget constraint under prioritization is obtained by pivoting the net neutrality constraint around $(X_0, X_1) = \left(\frac{1}{2}(\mu-1), \frac{1}{2}(\mu-1)\right)$, starting with a slope of $\frac{\omega_1^N}{\omega_0^N} = 1$ to attain slope $\frac{\omega_1^P}{\omega_0^P} = \mu$. Under priority, an increase in capacity μ thus brings both income and substitution effects, while it only brings an income effect under neutrality.

Letting $\left(X_0^i(\mu), X_1^i(\mu)\right)$, $i = N, P$ be the optimal solution under i, we obtain:

$$\begin{array}{ll} X_0^N(\mu) = \frac{\alpha}{\alpha+\beta}(\mu-1) & X_1^N(\mu) = \frac{\beta}{\alpha+\beta}(\mu-1) \\ X_0^P(\mu) = \frac{1}{2}\frac{\alpha}{\alpha+\beta}\left(\mu^2-1\right) & X_1^P(\mu) = \frac{1}{2}\frac{\beta}{\alpha+\beta}\frac{\mu^2-1}{\mu} \end{array} \tag{8}$$

From these demand functions, we can interpret a change from neutrality to priority as a simultaneous increase of "income" from $\mu - 1$ to $\frac{1}{2}\left(\mu^2 - 1\right)$ and of the non-prioritized "content" price from 1 to μ. We accordingly decompose the impact of a change in μ in an income effect and a substitution effect.

Note that if $\alpha \geq \beta$, consumption of prioritized content is greater than $\frac{1}{2}(\mu - 1)$ under neutrality. As prioritization makes the budget constraint pivot around $\frac{1}{2}(\mu - 1)$ while it reduces the relative price of the prioritized content, $\left(\frac{1}{2}X^N(\mu), \frac{1}{2}X^N(\mu)\right)$ is feasible under prioritization. Consequently, by a revealed preference argument, consumers prefer the prioritization regime. Prioritization involves a trade-off between a slower non-prioritized content and greater capacity only when consumers initially give more weight on the non-prioritized content. For this reason, hereafter we assume that $\alpha < \beta$.

We define the indirect utility function in regime i as $V^i(\mu) \equiv U^i\left(X_0^i(\mu), X_1^i(\mu)\right)$. From (6) and (8), we obtain:

$$V^N(\mu) = v(\mu - 1)^{\alpha+\beta} \tag{9}$$

$$V^P(\mu) = \left(\frac{1}{2}\right)^{\alpha+\beta} v\left(\mu^2 - 1\right)^{\alpha+\beta} \mu^{-\beta} \tag{10}$$

where $v \equiv \left(\frac{\alpha}{\alpha+\beta}\right)^{\alpha}\left(\frac{\beta}{\alpha+\beta}\right)^{\beta}$.

2.4 Market Functioning and Regulation

The ISP sells broadband connection to covered users at a fixed charge p. Users will agree to subscribe if their net utility $V^i(\mu) - p$ is larger than their outside option, that we normalize to zero. Hence, the ISP can extract all the surplus from the consumers and $p^i = V^i(\mu)$.

Under net neutrality, the ISP does not have financial relationships with the CP. The per location revenue is thus $R^N(\mu) \equiv V^N(\mu)$.

Under prioritization, the ISP gives an advantage to the prioritized content. Consequently, the traffic of the prioritized content increases by $X_0^P - X_0^N$ per location. We assume that content providers have no bargaining power for the implementation conditions of P management, so that the ISP is able to extract $a\left(X_0^P - X_0^N\right)$ per location from prioritization.[13] Under prioritization, the ISP per location revenue is then:

$$R^P(\mu) \equiv V^P(\mu) + a\left(X_0^P(\mu) - X_0^N(\mu)\right)$$

[13] We relax this assumption in Section 6.2.

A benevolent regulator monitors this two-sided market. Depending on the regulatory framework considered, it can determine the traffic management practice N or P, or the market coverage n, or both. The per location social benefit functions that she enters in her welfare function are given by:

$$B^i(\mu) \equiv V^i(\mu) + aX^i(\mu), \; i = N, P$$

If n markets are covered, the ISP profit is given by $\Pi^i(n,\mu) \equiv nR^i(\mu) - C(n,\mu)$, while social welfare is given by $W^i(n,\mu) = nB^i(\mu) - C(n,\mu)$.

Our analysis consists in comparing the performance of three regulatory frameworks. Under traffic management regulation (TMR), the regulator determines whether the ISP operates under net neutrality or prioritization, while the ISP chooses market coverage. Conversely, under universal service obligations (USO), the regulator imposes the market coverage and the ISP chooses the traffic management practice. Finally, under full regulation (FR), the regulator imposes both the traffic management practice and market coverage. We gauge the performance of these regulatory frameworks in terms of coverage and social welfare against the benchmark cases of a unregulated market (UM), where the ISP chooses both the traffic management regime and market coverage in order to maximize profit, and the first-best outcome (FB), where the welfare-maximizing regime and coverage are considered notwithstanding any market or institutional constraint.

3 Preliminary Results

In this section, we develop the fundamental results of the model that will lie behind the analysis of the ISP's and regulator's choices. We first make a comparison of the traffic management practices in terms of ISP revenue and social benefit and then deduce optimal coverages and management techniques for benchmark cases of unregulated markets and first-best.

3.1 Net Neutrality vs Priority: Comparative Statics

As a first step, we compare market outcomes obtained under neutrality and priority for given μ and n. Because costs are independent of the traffic management regime, we can abstract from them, so that outcome comparisons are made in terms of per location traffic, ISP revenue, and social benefit. Hereafter, for any function $F^i(\mu)$, $i = N, P$, we let $\Delta F \equiv F^P(\mu) - F^N(\mu)$.

For traffic, note that even though the change from neutrality to priority brings a positive income effect for both types of contents, the increased delay on the non-prioritized content can make consumers reduce total consumption. The next Lemma presents the threshold bandwidth for which prioritization increases total content consumption and utility.[14]

Lemma 1 *If $\alpha < \beta$,*

(a) $\Delta X \geq 0$ if and only if $\mu \gtreqless \mu_X \equiv \frac{\beta}{\alpha}$

(b) There exists a $\mu_V > \mu_X > 1$ such that $V^P(\mu) \geq V^N(\mu)$ if and only if $\mu \geq \mu_V$.

In contrast to Economides and Hermalin [15], where consumers' utility increases if and only if total consumption is increased, if $\mu \in (\mu_X, \mu_V)$, utility under priority is less than utility under neutrality even though total consumption is higher under priority. The difference comes from the fact that transmission speeds in Economides and Hermalin [15] are decision variables, so that they are set independently of capacity, instead of being determined by the M/M/1 queue, which introduces an interdependence of transmission speeds with capacity under priority.[15] Combined with the assumption that utility is additively separable in contents, content demands are independent in Economides and Hermalin [15], so that an increase of bandwitdh "is similar to more total income in a conventional consumer-choice model".[16] This income effect is also present in our model under both neutrality and priority,[17] but the interdependence of transmission speeds and capacity under priority adds a substitution effect. As a result, the change from neutrality to priority can involve both an increase in total consumption and a decrease in utility if the substitution effect, absent in Economides and Hermalin [15], is such that the utility loss associated to the decrease of the non-prioritized content consumption is not compensated by the increase in total consumption.

The next Proposition shows that prioritization gains a comparative advantage over neutrality as bandwidth is increased. However, the exact threshold for which prioritization dominates neutrality

[14] Again, because prioritization dominates neutrality in all respects whenever $\alpha \geq \beta$, we focus on the case where $\alpha \leq \beta$. The result nevertheless holds for $\alpha \geq \beta$, as we then obtain a threshold $\mu_X < 1$, so that $\mu > \mu_X$, $\forall \mu > 1$, meaning that total consumption is increased under priority whatever is μ. We would also obtain $\mu_V = 1$, meaning that $V^P > V^N$, $\forall \mu > 1$.

[15] See equation (2).

[16] Economides and Hermalin [15], p. 609.

[17] The Cobb-Douglas utility function can be considered as additively separable by taking its log form. Moreover, in both models, transmission speeds are by definition the same for all contents under neutrality.

depends on what is measured. Since $\Delta X_0 > 0$, the ISP has always an additional income from the CPs but this might be insufficient to compensate the lower revenue from the consumers. The ISP prefers prioritization when $\Delta R = \Delta V + a\Delta X_0 > 0$. This threshold from which priority increases the ISP revenue is less than the one that increases welfare. Indeed, since $\Delta X_1 < 0$ and $\Delta B = \Delta R + a\Delta X_1 < \Delta R$, it takes a greater bandwidth to make priority improve welfare than to make it improve the ISP revenue.

Considering that ΔB is equivalently equal to $\Delta V + \Delta X$, that $\Delta V\left(\mu_X\right) < 0$ and $\Delta X\left(\mu_V\right) > 0$, the minimum bandwidth necessary to obtain a social benefit increase is lower than the one necessary for obtaining an indirect utility increase but greater than the one necessary to obtain a traffic increase.

Proposition 1 *There exist a μ_R and a $\mu_B > \mu_X$ such that $\mu_R < \mu_B < \mu_V$ and*

1. $0 \geq \Delta R \geq \Delta B$, $\forall \mu \leq \mu_R$
2. $\Delta R > 0 \geq \Delta B$ *for* $\mu_R < \mu \leq \mu_B$
3. $\Delta R > \Delta B > 0$, $\forall \mu > \mu_B$

Proposition 1 implies that, for a given μ and n, if the ISP prefers net neutrality, then the regulator also prefers net neutrality. If the regulator prefers prioritization, then the ISP also prefers prioritization. More importantly, there exist cases where the regulator prefers net neutrality while the ISP prefers the priority regime. This is illustrated in Figure 1.

[Figure 1: Social welfare, ISP profit and consumer surplus differences]

The main message from the comparative statics is thus that the change from net neutrality to prioritization is the more likely the greater is the bandwidth. The regulator and the ISP however differ on the exact threshold for which they consider prioritization preferable to net neutrality.

3.2 Benchmark Coverages

In order to evaluate the performance of regulatory frameworks in the next section, we use two benchmarks: the first-best welfare maximizing benchmark and the unregulated market benchmark where the ISP maximizes its profit.

Unregulated Market Benchmark (UM) With no coverage regulation, the ISP maximizes its profit in either neutrality or prioritization:

$$\max_n \Pi^i(n,\mu) = nR^i(\mu) - C(n,\mu)$$

The first order condition is:

$$R^i(\mu) - C'_n(n,\mu) = 0 \tag{11}$$

Denoting the solution by n_I^i, we obtain:

$$n_I^i(\mu) = \frac{R^i(\mu)}{c\mu} \tag{12}$$

Since network deployment costs are independent of the traffic management practice, the practice that leads to the greater ISP coverage is the one that conveys the greater revenue. We then obtain the following result from Proposition 1.

Proposition 2 *As* $\mu \lesseqgtr \mu_R$,

$$\begin{aligned} n_I^P(\mu) &\lesseqgtr n_I^N(\mu) \\ \Pi^P\left(n_I^P(\mu),\mu\right) &\lesseqgtr \Pi^N\left(n_I^N(\mu),\mu\right) \end{aligned}$$

As a result of Proposition 2, the profit-maximizing coverage is $n_I(\mu) \equiv \max\left(n_I^N(\mu), n_I^P(\mu)\right)$ and the profit-maximizing regime is $\arg\max_i n_I^i(\mu)$.

First-Best Benchmark (FB) Under regime $i = N, P$, the first-best coverage is the solution to the following problem:

$$\max_n W^i(n,\mu) = nB^i(\mu) - C(n,\mu)$$

From first-order condition,

$$B^i(\mu) = C'_n(n,\mu)$$

Denoting the solution by n_*^i, we obtain :

$$n_*^i(\mu) = \frac{B^i(\mu)}{c\mu} \tag{13}$$

As $B^N(\mu) - R^N(\mu) = aX^N(\mu) > 0$ and $B^P(\mu) - R^P(\mu) = a\left(X_1^P(\mu) + X_0^N(\mu)\right) > 0$, social benefit is greater than ISP revenue under both management practices. As a result, the unregulated coverage is lower than the first-best coverage for a given μ. This fact and Proposition 1 then lead straightforwardly to the following Proposition.

Proposition 3 *(a) For $i = N, P$ and $\forall \mu$, $n_I^i(\mu) < n_*^i(\mu)$*

(b) As $\mu \lesseqqgtr \mu_B$,

$$\begin{aligned} n_*^P(\mu) &\lesseqqgtr n_*^N(\mu) \\ W^P\left(n_*^P(\mu), \mu\right) &\lesseqqgtr W^N\left(n_*^N(\mu), \mu\right) \end{aligned}$$

The first-best coverage is thus $n_*(\mu) \equiv \max\left(n_*^N(\mu), n_*^P(\mu)\right)$ and the first-best regime is $\arg\max_i n_*^i(\mu)$. Of course, the result that the first-best coverage is greater than the profit-maximizing coverage is the basic feature of any model on universal service. But in our model, the regulator can go counter not only to the ISP preferred coverage but also to its preferred traffic management practice. Interactions between coverages and regulatory frameworks are analyzed in the next section.

4 Choices of Traffic Management Practice and Market Coverage

In practice, the net neutrality debate and traffic management regulation have by and large been pursued independently of coverage considerations in general and on the presence or absence of universal service obligations in particular. We now analyze the interactions between the choice of the traffic management practice N or P and the choice of market coverage under different regulatory frameworks: the traffic management regulation (TMR), where the regulator chooses N or P but the coverage is chosen by the ISP, the universal service obligations (USO) where the regulator chooses the market coverage but not the traffic management practice, and the full regulation (FR) where the regulator chooses the coverage and the traffic management. Our objective is to identify the optimal regulatory framework and the cost of incomplete regulations.

4.1 Traffic Management Regulation (TMR)

Under TMR, the regulator can choose the traffic management practice N or P but cannot impose universal service obligations, so that market coverage is chosen by the ISP. Then, the fact that net neutrality has a comparative advantage for low bandwidth remains. However, since coverage is chosen by the ISP, for whom the comparative advantage of net neutrality vanishes at a lower level of bandwidth than for the regulator, the threshold capacity that makes the regulator prefer priority over net neutrality is lower than μ_B. Moreover, this threshold is also greater than μ_R since welfare is still greater under neutrality than under prioritization at $\mu = \mu_R$.

Proposition 4 *There exists a $\tilde{\mu}_0 \in (\mu_R, \mu_B)$ such that $W^P\left(n_I^P(\tilde{\mu}_0), \tilde{\mu}_0\right) \lesseqgtr W^N\left(n_I^N(\tilde{\mu}_0), \tilde{\mu}_0\right)$ as $\mu \lesseqgtr \tilde{\mu}_0$.*

Moreover, let n_T represent the coverage choice of the ISP under TMR. Then

(a) If $\mu \leq \mu_R$, the regulator chooses N and $n_T = n_I^N = n_I$

(b) If $\mu \in (\mu_R, \tilde{\mu}_0]$, the regulator chooses N and $n_T = n_I^N < n_I$

(c) If $\mu > \tilde{\mu}_0$, the regulator chooses P and $n_T = n_I^P = n_I$

Cases (a) and (c) are those where the choice of the regulator is aligned to the preferences of the ISP, so that TMR turns out to be irrelevant: welfare is the same than under UM since the ISP sets the unregulated market coverage anyway. The regulator makes a difference in case (b) where it imposes neutrality while the ISP would have chosen prioritization under UM. This makes the ISP choose a coverage that is lower than the one it would have chosen under UM. Rather surprisingly, regulation results in lower coverage and works to the detriment of unserved markets, to provide a higher utility in served markets.

Note also that if $\mu \in [\tilde{\mu}_0, \mu_B)$, which is a "sub-case" of (c), the regulator bends to the ISP preferred traffic management practice, priority in this case, even though it would have chosen neutrality if it were also in control of coverage. The per market consumer utility gain that neutrality would bring, which would justify its adoption in face of the welfare maximizing coverage, proves insufficient in face of the lower ISP coverage.

4.2 Universal Service Obligations (USO)

Under USO, the regulator can choose market coverage n_U, i.e. can impose universal service obligations, but does not have the power to determine the traffic management practice. We assume that ISP participation is not an issue in the sense that the ISP does not make a negative profit when the regulator imposes n_*^i, whatever is μ and $i = N, P$.[18] Then, independently of the coverage imposed by the regulator in the first stage, the ISP chooses P if and only if $\mu > \mu_R$, since priority leads to

[18] We relax this assumption in Section 5.1.

a higher revenue whatever is the coverage. As a result, for $\mu < \mu_R$ or $\mu > \mu_B$, the regime choice of the ISP corresponds to the one that the regulator favors and this allows the regulator to impose the first-best coverage. For $\mu \in (\mu_R, \mu_B)$, however, the choice is $n_U = n_*^P < n_*^I$ and the regulator is unable to attain first-best even though n_U is the welfare-maximizing coverage given the traffic management practice chosen by the ISP.

Proposition 5 *Let n_U be the welfare-maximizing coverage under USO regulation. Then*

(a) If $\mu \leq \mu_R$, the ISP chooses N and $n_U = n_*^N = n_*$

(b) If $\mu \in (\mu_R, \mu_B]$, the ISP chooses P and $n_U = n_*^P < n_*$

(c) If $\mu > \mu_B$, the ISP chooses P and $n_U = n_*^P = n_*$

Note that contrary to TMR, USO are always relevant, in the sense that they lead to an increase of welfare compared to an unregulated market whatever is the bandwidth level: even though the ISP and the regulator agree on the traffic management technique for $\mu \in (\mu_R, \mu_B]$, the regulator always wishes a greater coverage than the ISP does.

4.3 Full Regulation (FR)

Under FR, the regulator can impose both the traffic management practice and universal service obligations. Under the assumption that the ISP makes a non-negative profit at welfare-maximizing coverage, the regulator can attain the first-best if it imposes both the regulatory regime and universal service obligations. In this basic model, we thus assimilate full regulation to the FB benchmark. A distinction is introduced in section 5.1 with the conjunction of an ISP participation constraint and the impossibility for the regulator to freely make monetary transfers to the ISP.

4.4 Comparisons: Traffic Management and Market Coverage

In this section, we compare the coverage and the chosen traffic management practice under the three possible regulatory frameworks (FR, TMR, USO) as well as under UM. The comparison is summarized in Proposition 6 and illustrated in two figures. Before proceeding to these comparisons, we need a preliminary result showing that for sufficiently low bandwidths, net neutrality is clearly superior to prioritization as the ISP revenue under neutrality is not only greater than its revenue under prioritization, but is also greater than the social benefit under prioritization.

Proposition 6 *There exists a bandwidth threshold* $\tilde{\mu}_1 < \mu_R$ *such that* $B^P(\tilde{\mu}_1) \lesseqqgtr R^N(\tilde{\mu}_1)$ *and* $n_*^P \lesseqqgtr n_I^N$ *as* $\mu \lesseqqgtr \tilde{\mu}_1$*. Moreover:*

(a) If $\mu < \tilde{\mu}_1 < \mu_R$*, then* $n_U = n_*^N > n_T = n_I^N > n_*^P > n_I^P$ *and net neutrality is chosen under the four regulatory frameworks.*

(b) If $\tilde{\mu}_1 \leq \mu < \mu_R$*, then* $n_U = n_*^N > n_*^P \geq n_T = n_I^N > n_I^P$ *and net neutrality is chosen under the four regulatory frameworks.*

(c) If $\mu_R \leq \mu < \tilde{\mu}_0$*; then* $n_*^N > n_U = n_*^P > n_I^P \geq n_T = n_I^N$ *and net neutrality is chosen under FR and TMR, while prioritization is chosen under USO and UM.*

(d) If $\tilde{\mu}_0 \leq \mu < \mu_B$*, then* $n_*^N > n_U = n_*^P > n_T = n_I^P \geq n_I^N$ *net neutrality is chosen under FR, while prioritization is chosen under TMR, USO and UM.*

(e) If $\mu \geq \mu_B$*, then* $n_U = n_*^P \geq n_*^N > n_T = n_I^P > n_I^N$ *and Prioritization is chosen under four criteria.*

Figure 2 shows the choice of the traffic management practice N or P under the three regulatory frameworks FR, TMR, USO, and UM, respectively. For example $[N, N, P, P]$ means that prioritization is chosen in regulatory frameworks FR and TMR, while net neutrality is chosen under USO and UM.

[Figure 2: Net Neutrality or Prioritization : regulatory frameworks [FR,TMR,USO,UM]]

We see that, whatever is the regulatory framework, net neutrality has a comparative advantage for low bandwidths and priority, for high bandwidths. Moving from $\mu = 1$ to the right, there exists for each regulatory framework a bandwidth threshold from which priority becomes superior. When the traffic management practice is chosen by the regulator (under FR and TMR), net neutrality prevails for a greater bandwidth range than when it is chosen by the ISP, since neutrality allows the regulator to avoid the loss of fringe revenue that prioritization brings while this loss has no impact on the ISP. The switch to prioritization comes at a lower bandwidth under TMR than under FR because it is the ISP that chooses coverage based on the revenue function rather than on the social benefit function. In contrast, when traffic management is chosen by the ISP (in USO and UM), the switch to prioritization appears at a common bandwidth level because the coverage decision is already made at the ISP decision stage, so that the ISP chooses in a situation of a fixed cost and a per location revenue that is independent of coverage.

Figure 3 now illustrates the choice of coverage under each regulatory framework in function of bandwidth.

[Figure 3 : Proposition 6. n_T: in blue n_U: in red]

We see that coverage is always higher when it is chosen by the regulator (FR, USO) than by the ISP (TMR, UM). Note that, from (12) and (13), n_I^i and n_*^i are continuous functions of μ.[19] Discontinuities in the coverage paths n_U^i and n_T^i, represented by arrows in Figure 3, appear as decisions of traffic management practice and coverage are made by different agents: with USO, the switch to prioritization is made when profit are the same at μ_R while welfare-maximizing coverages chosen by the regulator would be the same at $\mu_B > \mu_R$; with TMR, the switch to prioritization is made at $\tilde{\mu}_0$ while profit-maximizing coverages chosen by the ISP would be the same at $\mu_R < \tilde{\mu}_0$. Again, this reflects the fact that, in cases where the regulator and the ISP favor different regimes, the regulator prefers neutrality while the ISP prefers prioritization.

From Propositions 3 and 5 as well as Figure 3, it is clear that the first best can be achieved with USO only when there is no disagreement between the regulator and the ISP on the preferred traffic management regime, that is either for low ($\mu \leq \mu_R$) bandwidth where they both prefer N or for high bandwidth ($\mu \geq \mu_B$) where they both prefer P. For those bandwidth values, regulating traffic management is useless and imposing USO is sufficient for having the first best. For the remaining intermediate values of μ, only full regulation can achieve the first best. For this parameter range, we discuss the cost of incomplete regulation i.e. the relative merits of TMR versus USO. The comparison is done in the following proposition.

Proposition 7 *There exists a $\tilde{\mu}_u \in (\mu_R, \tilde{\mu}_0)$ such that TMR leads to a higher welfare than USO if $\mu \in [\mu_R, \tilde{\mu}_u]$ and USO lead to a higher welfare than TMR for $\mu \geq \tilde{\mu}_u$.*

[19] We do not graph these functions as we cannot establish their exact shapes.

For $\mu \geq \tilde{\mu}_0$, TMR is useless as it replicates the unregulated market situation. Therefore for those parameters USO dominates TMR. For $\mu \in (\mu_R, \tilde{\mu}_0)$, USO allow the regulator to bring the welfare maximizing coverage given the P management practice chosen by the ISP. However, this management practice is not itself the welfare-maximizing one, so that the result is short of the first-best. Similarly, TMR allows to change the traffic management practice to N but at the cost of reducing market coverage. The optimal single-instrument policy of Proposition 7 trades-off these two dimensions.[20]

5 Extensions

5.1 ISP Participation Constraint

In our main analysis, we assumed that the ISP participation was not an issue when the regulator imposed $n_*(\mu)$. This is equivalent to assume there is no cost of public funds if the regulator has to subsidize the ISP for providing the USO coverage. Although this is in line with seminal papers on universal services, such as Anton et al. [1] and Valletti et al. [32], the question of the choice of the funding mechanism and its impact on the ISP behavior has quickly become a central theme in the literature.[21]

In this section, we take into account the possibility that the optimal USO coverage brings a deficit to the ISP so that there exists an ISP participation constraint that can be tight. We consider first the case where no compensation mechanism exists. This can be considered as one polar benchmark case, while our main model focused on another polar case of total compensation with no transaction cost. In conformity to the USO literature, we then focus on "self-funded" mechanisms where the ISP losses are "funded through cross subsidies or through taxes levied on consumers or firms involved in the market".[22] Note, however, that we analyze the case where funds

[20]Note that even in the range $(\mu_R, \tilde{\mu}_u)$, where welfare is higher under TMR than under USO, USO nevertheless bring a coverage n_*^P that is higher than the coverage n_I^N that is brought about by the regulator's choice of net neutrality under TMR.

[21]Seminal contributions on the subject are Chone et al. [13], [14].

[22]Chone et al [14], p. 1249. Note that models using direct subsidies can be considered as a special case of a model with a US fund where an exogenous shadow cost of public funds replaces the endogenous value of the Lagrange multiplier associated to the ISP participation constraint.

are levied from the CP side of the market, which is absent in the USO literature, rather than from consumers or from ISP competitors to the USO provider in an oligopolistic market.[23]

Assumptions about the capacity of the regulator to raise funds from CPs can easily abound. This is because tax avoidance from CPs is facilitated by the relative difficulty for governments to monitor CP activities for fiscal purposes. This difficulty is enhanced by the absence of geographical frontiers for data transmission.[24] However, the key point of this section is that, although the market coverage under USO is quantitatively modified by the ISP participation constraints and the presence of monetary transfers to the ISP, qualitative results are maintained: on the one hand, taking into account the participation constraint can reduce USO coverage compared to first-best coverage; on the other hand, the creation of a USO fund can counteract this effect through the increase of ISP revenue and, as we show that this increase is relatively greater for neutrality than for priority, the bandwidth threshold for which the ISP prefers priority under USO is then greater than μ_R.

Benchmark: No USO Fund We assume first that the regulator is unable to make any transfer to the ISP. In such a case, we must check whether its participation is ensured, i.e. whether there exists a range of bandwidth levels for which

$$\Pi^i\left(n_*^i(\mu),\mu\right) = n_*^i(\mu)\,R^i(\mu) - C\left(n_*^i(\mu),\mu\right) \geq 0 \tag{14}$$

From (5) and (13), this is equivalent to have:

$$R^i(\mu) \geq \frac{1}{2}B^i(\mu) \tag{15}$$

Since $\frac{\partial R^i}{\partial \mu} < \frac{\partial B^i}{\partial \mu}$ and B^i tends to infinity when μ tends to infinity, there exists a maximal bandwidth over which the ISP is not profitable. Since $\frac{\partial^2 R^i}{\partial a \partial \mu} < \frac{\partial^2 B^i}{\partial a \partial \mu}$, this maximal bandwidth decreases with a. The next lemma uses these facts to define the set M_*^i of bandwidths satisfying (14).

[23]In section 5.2, we adapt our analysis for the possibility of foreign ownership of the non-prioritized CPs, on the one hand, and the prioritized CPs, on the other hand.

[24]Accordingly, the OECD [24] observes, on the hand, that "because the digital economy is increasingly becoming the economy itself, it would not be feasible to ring-fence the digital economy from the rest of the economy for tax purposes", but on the other hand, that "certain business models and key features of the digital economy exacerbate base erosion and profit shifting risks."

Lemma 2 *There exists a* $\bar{\mu}_*^i(a)$ *such that* $\Pi^i\left(n_*^i(\bar{\mu}_*^i(a)), \bar{\mu}_*^i(a)\right) = 0$ *and* $M_*^i(a) = [1, \bar{\mu}_*^i(a)]$*. Moreover,* $\left(\bar{\mu}_*^i\right)'(a) < 0$.

It may seem counter-intuitive that an increase of ad price decreases the maximal feasible bandwidth of the network. The reason is that the first-best coverage increases as CPs become more profitable and, as coverage costs are convex, the first-best can prove to be too costly for the ISP. For bandwidth levels not in M_*^i, i.e. for μ such that $\Pi^i\left(n_*^i(\mu), \mu\right) < 0$, the choice of coverage under USO regulation is given by the ISP participation constraint. Coverage is thus the value n_π^i such that

$$\Pi^i\left(n_\pi^i(\mu), \mu\right) = n_\pi^i(\mu) R^i(\mu) - C\left(n_\pi^i(\mu), \mu\right) = 0$$

From (5) and (12), this implies $n_\pi^i(\mu) = 2n_I^i(\mu)$. If the participation constraint is binding under both N and P, the ISP will again choose the regime N under USO if and only if $\mu \leq \mu_R$, while the regulator will choose N under FR if and only if $\mu \leq \mu_B$. The participation constraint then introduces a difference between full regulation coverage and the first-best coverage. Since market coverage is chosen by the ISP under UM and TMR, results for these regulatory frameworks are not impacted by the participation constraint and results of Proposition 6 still hold. In the following proposition, we assume that $\bar{\mu}_*^i$ is so low that the participation constraint binds under USO and FR for all cases considered in Proposition 6.

Proposition 8 *If* $\mu > \max\{\bar{\mu}_*^N(a), \bar{\mu}_*^P(a)\}$*, and*

(a) If $\mu < \mu_R$*, then* $n_U = n_\pi^N > n_\pi^P > n_T = n_I^N > n_I^P$ *and net neutrality is chosen under the four regulatory frameworks*

(b) If $\mu_R \leq \mu < \tilde{\mu}_0$*; then* $n_U = n_\pi^P \geq n_\pi^N > n_I^P \geq n_T = n_I^N$ *and net neutrality is chosen under FR and TMR, while prioritization is chosen under USO and UM.*

(c) If $\tilde{\mu}_0 \leq \mu < \mu_B$*, then* $n_U = n_\pi^P > n_\pi^N > n_T = n_I^P > n_I^N$ *and net neutrality is chosen under FR, while prioritization is chosen under TMR, USO and UM.*

(d) If $\mu \geq \mu_B$*, then* $n_U = n_\pi^P > n_\pi^N > n_T = n_I^P > n_I^N$ *and prioritization is chosen under the four regulatory frameworks.*

If $\mu < \min\{\mu_*^N(a), \mu_*^P(a)\}$, the ISP participation constraint is not binding and we must turn back to the relevant case of Proposition 6 with respect to the exact value of μ. If $\min\{\mu_*^P(a), \mu_*^N(a)\} < \mu < \max\{\bar{\mu}_*^N(a), \bar{\mu}_*^N(a)\}$, the participation constraint binds under one and only one of the traffic management regime. Given the combinatorial nature of the possibilities and the fact that these possibilities do not involve new principles, we omit the presentation of results for such a case.[25]

USO Fund Assume now that the regulator is able to establish a US fund by seizing shares t_0 and t_1 of the prioritized and non-prioritized CPs' rents, respectively.[26] We let the shares be different for the two CPs as tax avoidance possibilities can in general differ across CPs.[27]

Collecting money on the CP side to finance infrastructure extension is an argument often used to justify the repeal of net neutrality and for allowing paid prioritization. A USO fund has the same purpose but it does not distort the consumers' demand for content.

The primary impact of the fund is to enlarge the set M_*^i for which the ISP is able to supply the first-best coverage. But it can also modify the ISP behavior. Hereafter, we consider μ in M_*^i given the existence of the fund and check whether the fund modifies results under USO and FR.

As the total tax proceeds are transferred to the ISP, the ISP revenues become:

$$\begin{aligned} R_t^N(\mu) &= V^N(\mu) + t_0 a X_0^N(\mu) + t_1 a X_1^N(\mu) > R^N(\mu) \\ R_t^P(\mu) &= V^P(\mu) + a(1-t_0)\Delta X_0 + t_0 a X_0^P(\mu) + t_1 a X_1^P > R^P(\mu) \end{aligned}$$

where, under prioritization, the ISP is able to directly charge the additional large CP profit, net of taxation. We thus have:

$$\Delta R_t = \Delta V + a\Delta X_0 + t_1 a \Delta X_f < \Delta R$$

[25]It can be shown that there is an ad price level a_μ such that $\bar{\mu}_*^P(a) \gtreqless \bar{\mu}_*^N(a)$ as $a \gtreqless a_\mu$.

[26]Alternatively, t_0 and t_1 can be considered as unit taxes on CPs. Moreover, this will make the analysis compatible with the extension on content diversity in section 6.3.

[27]For instance, it can be difficult to recover taxes from small CPs as their activities are difficult to monitor for the government – think for instance of bloggers or influencers. On their part, large CPs, even when they are domestic, can for instance practice tax shifting across countries. Fuchs [16] gives empirical evidence that some big digital companies intensively employ intangibles registered in low tax jurisdictions (as Ireland) and can operate in the market without necessarily being physically present. We come back below to the additional problem of foreign ownership. In section 6.3, the prioritized content will be supplied by a large CP, while the non-prioritized content will come from a fringe of small CPs.

Let n_t^i be the profit maximizing coverage under CP taxation. Since, from (12), market coverage is increasing with revenue, we immediately get that $n_t^i > n_I^i$. Moreover, since $\Delta R_t < \Delta R$, we have that the threshold bandwidth μ_t that is such that $\Delta R_t = 0$ is greater than μ_R. Prioritization becomes relatively less attractive than neutrality under USO funding because its introduction reduces funds from the fringe content without improving funding from the large content since the incremental transfer was already ensured without US fund. As a result, under USO regulation, prioritization is chosen for a lower range of bandwidths by the ISP than under UM. The range of disagreement between the regulator and the ISP on the preferred regime is reduced. There is also a discrepancy between the ISP choice of regime under UM and USO that did not exist without the US fund.

As usual, V and B are not modified by monetary transfers. Note that if the regulator is able to seize the totality of the CPs rent, i.e. if $t_0 = t_1 = 1$, then $R_t^i = B^i$, and the ISP espouses the regulator's preferences and simply maximizes welfare. As long as $t_1 < 1$, however, $\mu_t < \mu_B$.

In a nutshell, the possibility of establishing a US fund does not modify the main results of section 4, except for the fact that neutrality becomes favored by the ISP for a larger broadband range.

5.2 Foreign CPs

Although ISPs are generally regulated at the national level, content providers operate on a global market, as was quickly coined by the name *World Wide Web.* Apart from funding problems described in section 5.1, this raises the additional problem that some of the network value added is not considered in the welfare calculations of the regulator. In this section, we adapt the model for foreign ownership. While USO funding impacted on the ISP revenue without modifying the social benefit, foreign ownership impacts on the social benefit without impacting on the revenue.

Foreign non-prioritized CP Assume that the non-prioritized CP is a foreign firm. The regulator thus excludes its profit in its calculation of welfare, so that the social benefit becomes:

$$B_0^i(\mu) = B^i(\mu) - aX_1^i(\mu) = V^i(\mu) + aX_0^i(\mu) \tag{16}$$

The first-best coverage then becomes $n_0^i = \frac{B_0^i(\mu)}{c\mu}$. Since $B^i(\mu) > B_0^i(\mu) > R^i(\mu)$, it is clear that $n_*^i > n_0^i > n_I^i$. This is not surprising as welfare is lower under foreign ownership than under domestic ownership.

Note, however, that $\Delta B_0 = \Delta R$, so that disagreements over the choice of the regulatory regime disappear. The fact that prioritization reduces the non-prioritized CP profit, which the regulator took into account under domestic ownership, is ignored under foreign ownership. As a result, TMR becomes equivalent to UM and USO becomes equivalent to FR, the first-best coverage now being n_0^i. Letting $\hat{\mu} \in (\tilde{\mu}_1, \mu_R)$ be such that $B_0^P(\mu) \gtreqless R^N(\mu)$ as $\mu \gtreqless \hat{\mu}$, we obtain the following Proposition.

Proposition 9 *(a) If $\mu < \hat{\mu} < \mu_R$, then $n_U = n_0^N > n_T = n_I^N > n_0^P > n_I^P$ and net neutrality is chosen*

(b) If $\hat{\mu} \leq \mu < \mu_R$, then $n_U = n_0^N > n_0^P > n_T = n_I^N > n_I^P$ and net neutrality is chosen

(c) If $\mu \geq \mu_R$, then $n_U = n_0^P \geq n_0^N > n_T = n_I^P > n_I^N$ and prioritization is chosen.

Note that since the social losses due to non-prioritized CPs are not taken into account in the choice of the traffic management regime, prioritization is implemented for lower bandwidth levels than in our main case (i.e. $\mu_R \leq \tilde{\mu}_0$).

More importantly, Proposition 9 implies that TMR is useless when non-prioritized firms are foreign-owned. As regulator preference and ISP incentives are aligned in terms of management regime, the regulator can focus on USO. However, this case has limited scope because non-prioritized CPs are more likely of local interest.

Foreign Prioritized CP We now turn to the case where the prioritized CP is the foreign firm.[28] Then the social benefit becomes:

$$\begin{aligned} B_1^N(\mu) &= V^N(\mu) + aX_1^N(\mu) = B^N(\mu) - aX_0^N(\mu) < B^N(\mu) \\ B_1^P(\mu) &= V^P(\mu) + a\Delta X_0(\mu) + aX_1^P(\mu) = B^P(\mu) - aX_0^N(\mu) < B^P(\mu) \end{aligned}$$

so that $\Delta B_1(\mu) = \Delta B(\mu)$. Although social benefit is lower under foreign ownership than under domestic ownership, prioritization brings the same change in the social benefit for both cases. This is because the increased rent to the prioritized CP is seized whatever is the ownership structure. As the ISP revenue is not impacted by the prioritized CP ownership, bandwidth thresholds μ_R and

[28] This is a more likely case for any country except the United States.

μ_B stay the same, i.e. preferences of both the regulator and the ISP over the regulatory regimes are maintained.

However, since $B_1^i(\mu) < B^i(\mu)$, the first-best coverage n_f^i under foreign ownership of the prioritized CP is lower than n_*^i. The threshold $\tilde{\mu}_f$ for which $B_1^P(\mu)$ becomes greater than R^N is then greater than $\tilde{\mu}_1$. We also define $\hat{\mu}_f \in (\mu_R, \mu_B)$ as the bandwidth level such that $W_f^P\left(n_I^P, \hat{\mu}_f\right) = W_f^N\left(n_I^N, \hat{\mu}_f\right)$.

Proposition 10 *If the prioritized CP is a foreign firm, results in Proposition 6 applies by replacing accordingly $\tilde{\mu}_1$ by $\tilde{\mu}_f > \tilde{\mu}_1$, n_*^i by $n_f^i < n_*^i$ and $\tilde{\mu}_0$ by $\hat{\mu}_f$. Moreover $\hat{\mu}_f > \tilde{\mu}_0$.*

The last result in the Proposition 10 shows that since profits of prioritized CPs are not taken into account in the choice of the traffic management regime, net neutrality is implemented for higher bandwidth levels than in our main case.

6 Robustness Checks

6.1 Priority Capacity Endowment

For ease of presentation in section 2.2, we established a prioritization system under which waiting times were determined as if half the traffic observed under neutrality was given precedence in case of congestion. In this section, we show that our results are qualitatively similar for any system that gives priority to a share ρ that is greater than $\frac{\alpha}{\alpha+\beta}$, i.e. greater than the share of X_0^N in total traffic under net neutrality.

To understand this fact, assume that a share $\rho \in [\frac{\alpha}{\alpha+\beta}, 1)$ of traffic under net neutrality is given precedence under priority. Then the waiting times become:

$$\omega_0 = \frac{1}{\mu - \rho(\mu - 1)} = \frac{1}{(1-\rho)\mu + \rho} \tag{17}$$

$$\omega_1 = \mu\omega_0 \tag{18}$$

The bounds on ρ are set so that the priority service is able to accommodate X_0^N, so that a switch to priority service can be done without modifying consumption of the prioritized content. If ρ is less than $\frac{\alpha}{\alpha+\beta}$, it is impossible to prioritize content $X_0^N(\mu) = \frac{\alpha}{\alpha+\beta}(\mu - 1)$ at the quoted (minimum) speed $\frac{1}{\omega_0}$. Having $\rho > 1$ would result in a quoted speed $\frac{1}{\omega_0}$ greater than the net neutrality speed $\frac{1}{\bar{\omega}}$, which is in contradiction with the concept of prioritization.

Substituting waiting times (17) and (18) in (3) and taking X_0 as the numeraire gives the following prioritization "budget constraint":

$$X_0 + \mu X_1 = (\mu - 1) \cdot ((1 - \rho)\, \mu + \rho) \tag{19}$$

It is easily seen that constraint (4) is a particular case of (19) with $\rho = \frac{1}{2}$. Accordingly, (19) is geometrically the result of pivoting the net neutrality budget constraint around $(X_0, X_1) = (\rho\,(\mu - 1)\,, (1 - \rho)\,(\mu - 1))$, while constraint (4) follows the same procedure for $\rho = \frac{1}{2}$.

Note that the only difference in the two constraints lies in the RHS, so ρ impacts on the "income effect" of prioritization, but not on the substitution effect. Since the RHS is decreasing in ρ, the greater is ρ, the lower is the "income effect" of prioritization, so that V^P, R^P and B^P become decreasing functions with respect to ρ. Thus, qualitative results of our analysis are not modified, but quantitatively, the comparative advantage of prioritization becomes the greater the lower is ρ: whatever is the regulatory framework considered in Section 4, the range of bandwidths for which priority is the preferred regime is the larger the lower is ρ. Compared to our earlier analysis, taking $\rho > \frac{1}{2}$ would improve (quantitatively) the comparative advantage of neutrality, while taking $\rho < \frac{1}{2}$ gives a greater comparative advantage to prioritization.

At the lower bound $\rho = \frac{\alpha}{\alpha+\beta}$, prioritization implies a budget pivot around $\left(X_0^N(\mu)\,, X_1^N(\mu)\right)$, so that the optimal choice of contents under net neutrality is feasible under priority whatever is μ. In that case, prioritization has always a comparative advantage.[29] In the limit case $\rho = 1$, the budget pivot around corner point $(\mu - 1, 0)$ and makes any optimal choice under prioritization feasible under neutrality, so that neutrality has always the comparative advantage.

Apart from the ease of presentation, the choice of $\rho = \frac{1}{2}$ can be justified by an equity argument towards content providers. Note that $\left(\rho X_0^N(\mu)\,, (1 - \rho)\, X_1^N(\mu)\right)$ is the unique point that is feasible under both traffic management practices. So, $\rho = \frac{1}{2}$ is the unique value of ρ that ensures that in the case each content provider generates the same traffic under one management practice, this equal share of network is also feasible under the other regime. As both content providers generate the same revenue by unit of traffic and that this revenue is independent of the transmission time, so that it is unaffected by changes of transmission times if this is not accompanied by a change in

[29]Note, however, that this is no longer the case if we consider the impact of prioritization on content diversity as done in section 6.3. Then, the loss of diversity impacts negatively on utility and fringe content profit, so that the superiority of prioritization is not ensured.

traffic, both practices treat CPs equally when they are equally responsible for network usage and contribute equally to social benefit.

Of course, one can debate over such an interpretation of equity. The main point is that the analysis is not modified by any feasible choice of ρ.

6.2 ISP Bargaining Power

So far, we assumed that the ISP was able to fully capture the rent of the prioritized content provider. Even though this polar case of market power is somewhat in line with the assumption of a monopolistic ISP, it can be harder to justify in cases the prioritized content providers also have bargaining power. In this section, we show that our results in fact hold in all respects provided that the ISP has a "sufficiently large" bargaining power. We also show how these results are modified when this is not the case, i.e., when the bargaining power rather lies on the side of the content providers, as would be the case, for instance, if the prioritized content providers are part of the GAFAM group.

Let $\eta \in [0, 1]$ be an index of the ISP relative bargaining power *vis-à-vis* the prioritized CP, so that the revenue that the ISP can extract from the gain provided to the CP is now $R^P(\mu, \eta) = V + \eta a \Delta X_0$. Then, $\Delta R = \Delta V + \eta a \Delta X_0$. The lower is η, the more heavily the relative profitability of prioritization rests on the change of consumers' utility. Since the change of social benefit is independent of the ISP bargaining power, this means that the change of revenue brought by prioritization can now be lower than the change of the social benefit. As a result, the threshold bandwidth for which the ISP begins to prefer prioritization over net neutrality can be greater than the corresponding threshold for social benefit.

Lemma 3 *There exists a $\hat{\eta}$ such that $\mu_R(\eta) \lesseqqgtr \mu_B$ if and only if $\eta \gtreqqless \hat{\eta}$.*

For $\eta > \hat{\eta}$, the only impact of having $\eta < 1$ is that the range of bandwidths for which preferences of the ISP and the regulator over prioritization diverge is increased. This is because the ISP, by getting a lower part of industry revenue, takes into account a lower part of the social benefit in its evaluation of prioritization. Otherwise, results obtained in sections 2 to 4 are maintained once we substitute $\mu_R(\eta)$ to μ_R in lemmas and propositions.

Hereafter, we thus focus on the case $\eta \leq \hat{\eta}$. Then, the additional revenue obtained from the implementation of prioritization becomes too low for making the ISP evaluate more highly prioritization than the regulator, so that we now have $\Delta R \leq \Delta B$. Accordingly, we obtain the following proposition, where prioritization still eventually becomes the preferred choice as bandwidth is increased, but where it is now the ISP that is less prone to adopt prioritization.

Proposition 11 *If* $\eta \leq \hat{\eta}$*,*

1. $0 \geq \Delta B \geq \Delta R$, $\forall \mu \leq \mu_B(\eta)$
2. $\Delta B \geq 0 \geq \Delta R$ *for* $\mu_B \leq \mu \leq \mu_R(\eta)$
3. $\Delta B \geq \Delta R > 0$, $\forall \mu > \mu_R(\eta)$

As in Proposition 4, there still exist a $\tilde{\mu}_0$ such that $W^P\left(n_I^P(\tilde{\mu}_0), \tilde{\mu}_0\right) \lesseqqgtr W^N\left(n_I^N(\tilde{\mu}_0), \tilde{\mu}_0\right)$ as $\mu \lesseqqgtr \tilde{\mu}_0$ and a $\tilde{\mu}_1 < \mu_R$ such that $B^P(\tilde{\mu}_1) \lesseqqgtr R^N(\tilde{\mu}_1)$ as $\mu \lesseqqgtr \tilde{\mu}_1$, but now $\tilde{\mu}_1$ and $\tilde{\mu}_0$ such that $\tilde{\mu}_1 < \mu_B < \tilde{\mu}_0 < \mu_R(\eta)$. Results for the case $\eta \leq \hat{\eta}$ can be summed up in the following proposition, which echoes Proposition 6 that applied to the case $\eta > \hat{\eta}$.

Proposition 12 *If* $\eta < \hat{\eta}$ *and*

(a) *If* $\mu < \tilde{\mu}_1 < \mu_B$, $n_U = n_*^N > n_T = n_I^N > n_*^P > n_I^P$ *and net neutrality is chosen under the four regulatory frameworks*

(b) *If* $\tilde{\mu}_1 \leq \mu < \mu_B : n_U = n_*^N > n_*^P \geq n_T = n_I^N > n_I^P$ *and net neutrality is chosen under the four regulatory frameworks*

(c) *If* $\mu_B \leq \mu < \tilde{\mu}_0 : n_*^P > n_U = n_*^N \geq n_T = n_I^N > n_I^P$ *and prioritization is chosen under FR, while net neutrality is chosen under UM, TMR and USO*

(d) *If* $\tilde{\mu}_0 \leq \mu < \mu_R : n_*^P > n_U = n_*^N \geq n_I^N > n_T = n_I^P$ *and prioritization is chosen under FR and TMR, while neutrality is chosen under UM and USO*

(e) *If* $\mu \geq \mu_R : n_U = n_*^P > n_*^P \geq n_T = n_I^P > n_I^N$ *and prioritization is chosen under the four regulatory frameworks*

The results in Proposition 6 were established with the highest bargaining power for the ISP. This can be viewed as a favorable upper bound in terms of private coverages in the analysis. Of course, since the bargaining power influences only monetary transfers between the ISP and CPs, welfare is not impacted by it. But if prioritization is meant to help investment in networks, as is often advanced by its proponents, and if the choice of the traffic management practice is left to the ISP, as is the case for UM and USO regulations, the bargaining power of the ISP *vis-à-vis* CPs becomes an important factor to take into account.

Note that relaxing the assumption that the ISP has full market power does not modify the fact that the comparative advantage of net neutrality is the more likely the lower is the bandwidth.

6.3 Content Diversity

A primary benefit that is attributed to net neutrality is to promote network access for content providers, thus ensuring competition, innovation and diversity at this end of the market. Reggiani and Valletti [28] confirm this conjecture. In this section, we show that these considerations are easily integrated in our model.

Assume that the anonymous CPs we considered up to now are in fact belonging to two classes that are different in nature. The CP denoted 0 is a large CP (say Google or Facebook), while the CP formerly denoted 1 becomes a fringe of m small CPs (each denoted by j). CPs in the fringe face a fixed entry cost F so that their individual profits are given by $\Pi_j = anX_j - jF$. There is free entry in fringe content supply, so that Π_m will be nil at equilibrium and the number of fringe content types m is endogenous. We let $X_f \equiv \sum_{i=1}^{m} X_j$ be aggregate fringe traffic and $X = X_0 + X_f$ be the total traffic on the network.

Consumers now value diversity of the fringe contents. More precisely, the utility function has the following separable form in the large and fringe contents:

$$U\left(X_0, Z_f\left(X_1, ..., X_m\right)\right) = X_0^{\alpha} \cdot Z_f$$

where Z_f is a CES index of the overall fringe consumption that takes into account both substitutability among contents and the number of varieties:[30]

$$Z_f \equiv \left(\sum_{j=1}^{m} X_j^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\beta\sigma}{\sigma-1}} \qquad (20)$$

[30] This representation of preferences in a model of monopolistic competition is borrowed from Belleflamme and Peitz [2], p. 88.

In this expression, $\sigma > 1$ is the elasticity of substitution between any two fringe contents and $\beta < \frac{\sigma-1}{\sigma}$ is the degree of homogeneity of the CES function. The bound on β is set in order to ensure strict concavity of the index.

As the same transmission time applies to any fringe content X_j and the weight given on X_j in Z_f is the same for all j, total fringe consumption will be evenly distributed among fringe CPs, i.e. $X_j = \frac{X_f}{m}$. Substituting this value in (20) gives $Z_f = X_f^{\beta} m^{\frac{\beta}{\sigma-1}}$ and the consumer problem can be written as:

$$\begin{aligned} \max_{X_0, X_f} \quad & X_0^{\alpha} X_f^{\beta} m^{\delta} \\ \text{s.t.} \quad & \omega_0^i X_0 + \omega_f^i X_f = \mu - 1 \end{aligned}$$

where $\delta \equiv \frac{\beta}{\sigma-1}$ and where ω_0^i and ω_f^i are transmission times of large and fringe CPs, respectively, under regulatory regime i. Note that in this formulation, consumers value diversity per se. Since m is a constant in this problem, optimal solutions are still given by (8). Note the case of no fringe diversity that we considered up to now is a particular case where $m = 1$.

Equilibrium number of fringe CPs is obtained from the zero-profit condition. Recalling that $X_m = \frac{X_f}{m}$, we obtain:

$$m^i(n, \mu) = \left(\frac{anX_f^i(\mu)}{F} \right)^{\frac{1}{2}}, \ i = N, P \tag{21}$$

The fringe aggregate profit is then:

$$\Pi_f^i(n, \mu) = anX_f^i - \frac{(m^i)^2}{2} nF = \frac{1}{2} anX_f^i(\mu)$$

Whatever is the traffic management practice, both an increase in capacity and coverage favor diversity and there is complementarity between capacity and coverage with respect to diversity. However, since demand of fringe content is lower under prioritization than under neutrality, a shift from neutrality to priority lowers diversity of content. The difference tends to be attenuated when market coverage is increased because, as the fixed cost of fringe CPs is independent of coverage, the exit of firms because of lower demand becomes less severe as coverage is increased.

Proposition 13 *In both regimes N and P, $m'_\mu > 0$, $m'_n > 0$, $m''_{\mu n} > 0$. Moreover, letting $\Delta m = m^P(n, \mu) - m^N(n, \mu)$, we obtain $\Delta m < 0$, $\frac{\partial(\Delta m)}{\partial n} < 0$ and $\frac{\partial^2(\Delta m)}{\partial n^2} > 0$*

Per location utility function $V^i = U\left(X_0(\mu), X_f(\mu), m(n,\mu)\right)$ then depends on market coverage:

$$V^N(n,\mu) = n^{-\frac{1}{2}\delta} v (\mu - 1)^{\alpha+\beta+\frac{1}{2}\delta} \tag{22}$$

$$V^P(n,\mu) = \left(\frac{1}{2}\right)^{\alpha+\beta} n^{-\frac{1}{2}\delta} v \left(\mu^2 - 1\right)^{\alpha+\beta+\frac{1}{2}\delta} \mu^{-\left(\beta+\frac{1}{2}\delta\right)} \tag{23}$$

where $v \equiv \left(\frac{a}{f}\right)^{\frac{1}{2}\delta} \left(\frac{\alpha}{\alpha+\beta}\right)^{\alpha} \left(\frac{\beta}{\alpha+\beta}\right)^{\gamma\beta+\frac{1}{2}\delta}$. Note that per location indirect utility is convex with respect to n, but aggregate utility nV^i, which enter the objective functions of the regulator and the ISP, are concave in n. The functional forms of (22) and (23) with respect to μ are identical to those of (9) and (10). Results for this version of the model are thus qualitatively similar to those obtained with those in sections 3 and 4. However, as fringe content becomes more valuable with diversity while neutrality is the regime that is the most favorable to diversity, the ranges of intervals for which prioritization is chosen in proposition 6 are simply reduced.

In summary, taking into account diversity introduces a positive network externality from non-prioritized content demand: the greater revenue that induces a greater fringe content demand helps support more variety, which is retroactively valued as such. As both coverage and capacity induces more fringe content consumption, they both contribute to this positive network externality. These impacts are, however, less important for priority. As a result, both the ISP and the regulatory agency recur less often to prioritization when diversity of content is taken into account.

Note that our results on diversity are in line with those of Reggiani and Valletti [28] in particular and with the general argument in favor of net neutrality that was first stated by Wu [34].

7 Conclusion

We have integrated the analysis of net neutrality and universal service obligations in a single model. We have shown that the comparative advantage of prioritization over net neutrality for extending broadband market coverage is positively correlated with bandwidth or, in other words, with the data transmission capacity of the network. The reason is that the greater is the network capacity, the greater is the relative gain of prioritization in terms of total traffic carried against the utility loss associated to non-prioritized content displacement. This correlation is robust to changes of our basic assumptions that CPs are domestically owned and that there is no cost of public funds in transfers between the regulator and the ISP: only the exact bandwidth thresholds for which the ISP or the regulator preferences switch from neutrality to prioritization are sensitive to these assumptions. Accordingly, prioritization, or equivalently, the repeal of net neutrality, looks relatively less attractive for the ISP if it benefits revenue from a USO fund or if it has bargaining power *vis-à-vis* content providers. It also looks relatively less attractive for both the ISP and the regulator when non-prioritized content is diversified. It is relatively more attractive for the regulator when the non-prioritized CPs are foreign owned. Finally, foreign ownership of prioritized CPs does not modify the comparative advantage of one management technique over the other.

We have considered three regulatory frameworks that represent the different combinations of free or regulated traffic management and market coverage:

- Traffic management regulation (TMR), where the regulator decides whether net neutrality is maintained or repealed while the ISP chooses the market coverage
- Universal service obligations (USO), where the regulator imposes the market coverage to the ISP while the latter chooses the traffic management technique
- Full regulation (FR), where both traffic management and coverage are set by the regulator

Welfare obtained under these regulatory frameworks were compared to unregulated market (UM) and first-best (FB) benchmarks.

Whatever is the bandwidth level, TMR fails to maximize welfare because the monopolistic profit maximizing coverage falls short of the first-best coverage. TMR can nevertheless improve on the UM coverage when the bandwidth level stands in a range of intermediate values for which there is a conflict between the profit-maximizing and the first-best traffic management technique. In this range, the regulator proves less prompt to repeal net neutrality than the ISP, so that prioritization can be forbidden even though the ISP would adopt it. For relatively low or relatively high bandwidth levels, imposing TMR is useless as interests of the ISP and the regulator converge for net neutrality in case of narrow bandwidth, and for priority in case of large bandwidth.

On the contrary, USO always improve welfare compared to an unregulated market and even if they do not constitute full regulation, they can be welfare-maximizing in presence of sufficiently low or sufficiently high bandwidth. This corresponds again to bandwidth ranges where there is agreement between the regulator and the ISP. USO miss the first-best outcome for intermediate levels because they can let the ISP prioritize traffic while net neutrality should have been maintained.

So, in a range of intermediate values of bandwidth, full regulation is required to attain maximum welfare because USO cannot impose net neutrality. If only one regulation is to be imposed, then TMR proves to be welfare-superior to USO for the lower values of this range, while the reverse is true for the higher values. The trade-off between TMR and USO involves a trade-off between optimal traffic management and optimal coverage: for lower bandwidth levels, the welfare loss of lower coverage under TMR is less than the welfare loss of the net neutrality repeal under USO, and conversely for upper bandwidth levels. But globally, universal service obligations appear to be a stronger regulatory instrument than the imposition of the traffic management method since USO reach first-best for low and high bandwidth levels, while TMR never allows to reach maximal welfare, and it is only for a subset of intermediate values where USO fail to reach first-best that TMR is welfare-superior. It is important to note, however, that if USO lead to the first-best for high bandwidth, it is because our definition of USO leaves the ISP chooses the traffic regime. In practice, in most countries to the notable exception of US, this requires that regulators repeal net neutrality.

The main policy implication is thus that net neutrality should eventually be repealed in face of the ever-increasing bandwidth requirements of internet applications and contents. As video streaming and on-line video games are a decade-old phenomenon, our model seems in line with the history of internet. At early days of narrowband internet, after some experiences of closed networks such as AOL, net neutrality became the dominant traffic management practice on the internet well before the term was coined by Wu ([34]). Net neutrality contributed to the universal adoption of internet and to the diversity of content that it delivered, with an important increase of content requiring broadband in the early 2010's. But, in a certain sense, it became victim of its own success as it fed debates on its economic efficiency as concerns over congestion and misallocation of traffic due to equal treatment of contents of different time sensitivities arose (Peitz and Schuett [25]). ISPs were first to ask for a repeal of net neutrality and the US became the first country to act in this sense in 2018.[31] If we consider that bandwidth requirement is currently in an intermediate stage of growth, our model suggests that there will be a tendency to repeal net neutrality.

[31] Since large CPs that are likely to be prioritized are domestic firms in US, the fact that US is the first country to repeal net neutrality is also in line with Proposition 10 that shows that net neutrality is implemented for higher bandwidth levels when prioritized CPs are foreign-owned.

These results constitute a contribution to both the net neutrality (NN) and the USO literature. First, since investment in infrastructure in the NN literature generally considers a fixed number of end-users, they focus on the intensive margin. The choice of the market coverage adds a trade-off between the extensive margin and the intensive margin. This puts the debate on net neutrality into a better perspective as a change in extensive margin has more impact on the network benefits than one on the intensive margin. Second, while the USO literature studies “one-sided” markets, prioritization introduces a funding method for market expansion for a two-sided market. This relaxes the constraints on universal service financing.

In this first paper to integrate net neutrality and universal service, we have omitted some topics studied in either one or both literature strands. The most important limitation of our model is the fact that the ISP is monopolistic and CPs are price-takers. Future work could be inspired by the treatment of duopolistic ISPs in the USO literature (Valletti et al. [32], for instance) and the analysis, in the NN literature, of CPs able to invest in their own infrastructure to improve their quality of service (Choi and Kim [10], for instance).

References

[1] Anton, J., Vander Weide, J. H. and N. Vettas, 2002. “Entry Auctions and Strategic Behavior under Cross-Market Price Constraints”, *International Journal of Industrial Organization*, 20, 611-629.

[2] Belleflamme, P. and M. Peitz, 2015. *Industrial Organization,* Cambridge University Press.

[3] Bennet, M. F., King, A. S. Jr, Machin III, J. and R. Porter: “Open Letter”, available through https://9to5mac.com/2021/03/04/us-high-broadband/, accessed October 7th, 2021.

[4] Bourreau, M., F. Kourandi and T. Valletti, 2015. “Net Neutrality with Competing Internet Platforms”, *The Journal of Industrial Economics,* LXIII, 30-73.

[5] Bourreau, M., Feasy, R. and S. Hoernig, 2017. “Demand-Side Policies to Accelerate the Transition to Ultrafast Broadband”, Centre on Regulation in Europe (CERRE).

[6] BroadbandNow, “https://broadbandnow.com/report/fcc-broadband-definition/”

[7] Calzada, J. and M. Tselekounis, 2018. "Net Neutrality in an Hyperlinked Internet Economy", *International Journal of Industrial Organization,* 59, 190-221.

[8] Canadian Radio-television and Telecommunications Commissions, 2021. "Broadband Fund: Closing the digital divide in Canada", https://crtc.gc.ca/eng/internet/internet.htm, accessed October 7^{th}, 2021.

[9] Cornière, A. et G. Taylor, 2014. "Integration and Search Engine Bias", *Rand Journal of Economics*, 45, 576-597.

[10] Choi, J. P., D.-S. Jeon and B.-C. Kim, 2018. "Net Neutrality, Network Capacity, and Innovation at the Edges", *The Journal of Industrial Economics,* LXVI, 172-204.

[11] Choi, J. P., Jeon, D.-S. et Kim, B.-C. (2015). Net Neutrality, Business Models, and Internet Interconnection. American Economic Journal: Microeconomics, 7, 104–141.

[12] Choi, J.P, B.-C. Kim, 2010. "Net Neutrality and Investment Incentives", *Rand Journal of Economics,* 41, 446-471.

[13] Chone, P., Flochel, L. and A. Perrot (2000), "Universal Service Obligations and Competition", *Information Economics and Policy*, 12, 249-259.

[14] Chone, P., Flochel, L. and A. Perrot (2002), "Allocating and Funding Universal Service Obligations in a Competitive Market", *International Journal of Industrial Organization,* 20, 1247-1276.

[15] Economides, N. and B. E. Hermalin, 2012. "The Economics of Network Neutrality", *Rand Journal of Economics,* 43, 602-629.

[16] Fuchs, C. (2018) "Google and Facebook's Tax Avoidance Strategies", Chapter 4, in *The Online Advertising Tax as the Foundation of a Public Service Internet,* CAMRI Extended Policy Report, University of Westminster Press.

[17] Gautier, A. and X. Wauthy, 2010. "Price Competition under Universal Service Obligations", *International Journal of Economic Theory*, 6, 311–326.

[18] Garcia-Calvo, A., 2012. “Universal Service Policies in the Context of National Broadband Plans”, *OECD Digital Economy Papers*, No. 203, http://dx.doi.org/10.1787/5k94gz19flq4-en.

[19] Hayel, Y., B. Tuffin, “Pricing for Heterogeneous Services at a Discriminatory Processor Sharing Queue”, 2005. In: *International Conference on Research in Networking,* Springer, 816-827.

[20] Katz, M. L., 2017. “Whither U.S. Net Neutrality Regulation?” *Review of Industrial Organization*, 50, 441-468.

[21] Krämer, J., Wiewiorra, L. and C. Weinhardt, 2013. “Net Neutrality: A Progress Report”, *Telecommunications Policy*, 37, 794-813.

[22] Kramer, J., L Wiewiorra, 2012. “Network Neutrality and Congestion Sensitive Content Providers”, *Information Systems Research*, 23, 1303-1321.

[23] McMenemy, D., 2022. “Internet Access and Bridging the Digital Divide: The Crucial Role of Universal Service Obligations in Telecom Policy. In: Smiths, M. (ed.), *Information for a Better World: Shaping the Global Future.* Series: Lecture Notes in Computer Science, 13192. Springer: Cham. pp. 122-134.

[24] OCDE, 2015. *Addressing the Tax Challenges of the Digital Economy*, Action 1 - 2015 Final Report, OECD/G20 Base Erosion and Profit Shifting Project, Éditions OCDE, Paris, https://doi.org/10.1787/9789264241046-en, accessed November 19^{th}, 2021.

[25] Peitz, M. and F. Schuett, 2016. “Net Neutrality and Inflation of Traffic”, *International Journal of Industrial Organization*, 46, 16-62.

[26] Poudou, J.-C. and M. Roland, 2017. “Equity Justifications for Universal Service Obligations”, *International Journal of Industrial Organization*, 52, 63-95.

[27] Poudou, J.-C. and M. Roland, 2014. “Efficiency of Uniform Pricing in Universal Service Obligations”, *International Journal of Industrial Organization*, 37, 141-152.

[28] Reggiani, C. and T. Valletti, 2016, “Net Neutrality and Innovation at the Core and at the Edge”, *International Journal of Industrial Organization,* 45, 16-27.

[29] Rosenworcel, J., 2016, "Bringing the Connected Future to All Americans, May 11, 2012 – January 3, 2017", https://www.fcc.gov/news-events/blog/2016/12/30/bringing-connected-future-all-americans-may-11-2012-january-3-2017, accessed October 7th, 2021.

[30] Sacoto-Cabrera, E. J., Guijarro, L., Vidal, J. R. and V. Pla, 2020. "Economic Viability of Virtual Operators in 5G via Network Slicing", *Future Generation Computer Systems,* 109, 172-187.

[31] Spulber, D. F. and C. Yoo, "Networks in Telecommunications", Cambridge University Press, 2009.

[32] Valletti, T., Hoernig, S. and P. Barros, 2002. "Universal Service and Entry: The Role of Uniform Pricing and Coverage Constraints", *Journal of Regulatory Economics*, 21, 169-190.

[33] Verizon, 2021. "Bandwidth", https://www.verizon.com/info/definitions/bandwidth/, accessed October 7^{th}, 2021.

[34] Wu, T., "Network Neutrality, Broadband Discrimination", *Journal of Telecommunications and High Technology,* 2, 141-179.

Appendix

Proof of Lemma 1.

(a) From (8),

$$\Delta X \equiv X^P(\mu) - X^N(\mu) = \frac{(\mu-1)^2}{2(\alpha+\beta)\mu}(\alpha\mu-\beta) \lesseqqgtr 0 \text{ as } \mu \lesseqqgtr \mu_X \equiv \frac{\beta}{\alpha}$$

(b) From (9) and (10),

$$\frac{V^P}{V^N} = \left(\frac{1}{2}\right)^{\alpha+\beta}(\mu+1)^{\alpha+\beta}\mu^{-\beta}$$

so that

$$V^P \lesseqqgtr V^N \text{ as } G(\mu) \equiv \frac{\ln\left(\frac{\mu+1}{2}\right)}{\ln\mu - \ln\left(\frac{\mu+1}{2}\right)} \lesseqqgtr \mu_X \tag{24}$$

As $G(\mu)$ is an strictly increasing function of μ such that $\lim_{\mu\to1} G(\mu) = 1$ and $\lim_{\mu\to\infty} G(\mu) = \infty$, then it exists $\mu_V : G(\mu_V) = \mu_X > 1$. Note that

$$1 \leq G(\mu) \leq \mu \text{ for } \mu \geq 1 \tag{25}$$

so $\mu_V > G(\mu_V) = \mu_X$. Moreover $\mu - G(\mu)$ is an strictly increasing concave function of μ.

Proof of Proposition 1. Note that for all $\mu > 0$, $\Delta B(\mu) = \Delta V(\mu) + a\Delta X_0(\mu) + a\Delta X_1(\mu) < \Delta V(\mu) + a\Delta X_0(\mu) = \Delta R(\mu)$ since $\Delta X_1(\mu) < 0$. For the large-biased case ($\alpha \geq 0.5$), since $\Delta V(\mu) \geq 0$ and $\Delta X(\mu) \geq 0$, $\forall \mu \geq 1$, we have $\mu_R = \mu_B = 1$. Consider now the fringe biased case ($\alpha \leq 0.5$). Since $\Delta V(\mu) \lesseqqgtr 0$ as $\mu \lesseqqgtr \mu_V$, $a\Delta X_0(1) = 0$, $a \lim_{\mu\to\infty} \Delta X_0(\mu) \to \infty$ and $\frac{\partial(\Delta X_0(\mu))}{\partial\mu} > 0$, $\forall\mu$, there exists a $\mu_R < \mu_V$ such that $\Delta R(\mu) = \Delta V(\mu) + a\Delta X_0(\mu) \lesseqqgtr 0$ as $\mu \lesseqqgtr \mu_R$.

Similarly, since $\Delta R \lesseqqgtr 0$ as $\mu \lesseqqgtr \mu_R$, $a(\Delta X(1)) = 0$, $\lim_{\mu\to\infty} a(\Delta X(\mu)) \to \infty$, and $\Delta B < 0 = \Delta R$ at $\mu = \mu_R$, there exists a $\mu_B > \mu_R$ such that $\Delta B \lesseqqgtr 0$ as $\mu \lesseqqgtr \mu_B$. As $\mu_X < \mu_V$, $\Delta B(\mu_X) = \Delta V(\mu_X) < 0$, so that $\mu_B > \mu_X$, and $\Delta B(\mu_V) = \Delta X(\mu_V) > 0$, so that $\mu_B < \mu_V$.

Proof of Proposition 2. Assume that $\mu < \mu_R$. From Lemma 1, $\Delta R < 0$ so that $n_I^P < n_I^N$. $\Delta R < 0$ also implies that $\Pi^P(n_I^P, \mu) < \Pi^N(n_I^P, \mu)$; since profits $\Pi^i(n, \mu)$ are strictly concave in n, n_I^N is a unique maximum to Π^N and $\Pi^N(n_I^P, \mu) < \Pi^N(n_I^N, \mu)$. We thus have $\Pi^P(n_I^P, \mu) < \Pi^N(n_I^N, \mu)$. The proof is similar for $\mu \geq \mu_R$.

Proof of Proposition 3. As $B^i(\mu) > R^i(\mu)$ for all μ, by definitions (12) and (13) of coverages, we have the first result. Now, consider the case where $\mu < \mu_B$. From Lemma 1, $\Delta B < 0$, so that $n_*^P < n_*^N$. Moreover, $\Delta B < 0$ also implies that $W^P(n_*^P, \mu) < W^N(n_*^P, \mu)$; since welfare $W^i(n, \mu)$ is strictly concave in n, n_*^N is a unique maximum to W^N and $W^N(n_*^P, \mu) < W^N(n_*^N, \mu)$. We thus have $W^P(n_*^P, \mu) < W^N(n_*^N, \mu)$. The proof is similar for $\mu \geq \mu_B$.

Proof of Proposition 4. Note that for all μ, coverage solutions are given by $N(x, \mu) = \frac{x}{c\mu}$, which is an increasing function in x. So we have:

$$\begin{aligned} & W^P\left(n_I^P(\mu), \mu\right) - W^N\left(n_I^N(\mu), \mu\right) \\ = \; & N\left(R^P(\mu), \mu\right) B^P(\mu) - C\left(n_I^P(\mu), \mu\right) - N\left(R^N(\mu), \mu\right) B^N(\mu) + C\left(n_I^N(\mu), \mu\right) \end{aligned}$$

At $\mu = \mu_R$, we have:

$$\begin{aligned} & W^P\left(n_I^P(\mu_R), \mu_R\right) - W^N\left(n_I^N(\mu_R), \mu_R\right) \\ < \; & \left[N\left(R^P(\mu_R), \mu_R\right) - N\left(R^N(\mu_R), \mu_R\right)\right] B^N(\mu_R) - C\left(n_I^P(\mu_R), \mu_R\right) + C\left(n_I^N(\mu_R), \mu_R\right) = 0 \end{aligned}$$

where the inequality comes from the fact that $B^P(\mu_R) < B^N(\mu_R)$ and the equality, from the fact that $R^P(\mu_R) = R^N(\mu_R)$ and $n_I^P(\mu_R) = n_I^N(\mu_R)$. Similarly, at $\mu = \mu_B$, we have:

$$\begin{aligned} & W^P\left(n_I^P(\mu_B), \mu_B\right) - W^N\left(n_I^N(\mu_B), \mu_B\right) \\ > \quad & N\left(R^N(\mu_B), \mu_B\right)\left(B^P(\mu_B) - B^N(\mu_B)\right) + C\left(n_I^N(\mu_B), \mu_B\right) > 0 \end{aligned}$$

where the first inequality comes from the fact that $R^P(\mu_B) > R^N(\mu_B)$ and the second, from the fact that $B^P(\mu_B) = B^N(\mu_B)$. By continuity, there exists a $\tilde{\mu}_0 \in (\mu_R, \mu_B)$ such that $W^P\left(n_I^N(\tilde{\mu}_0), \tilde{\mu}_0\right) = W^N\left(n_I^P(\tilde{\mu}_0), \tilde{\mu}_0\right)$. Moreover, from Propositions 1 and 2, for $\mu < \mu_R < \mu_B$, $n_*^N(\mu) > n_I^N(\mu) > n_I^P(\mu)$ and $W^P(n, \mu) - W^N(n, \mu) < 0$, so that we have

$$W^P\left(n_I^P(\mu), \mu\right) < W^N\left(n_I^P(\mu), \mu\right) < W^N\left(n_I^N(\mu), \mu\right) < W^N\left(n_*^N(\mu), \mu\right)$$

So this proves that $W^P\left(n_I^P(\mu), \mu\right) - W^N\left(n_I^N(\mu), \mu\right) < 0$ for $\mu < \mu_R$. Identically, for $\mu > \mu_B > \mu_R$, $n_I^P(\mu) > n_I^N(\mu)$ and $W^P(n, \mu) - W^N(n, \mu) > 0$ so

$$W^N\left(n_I^N(\mu), \mu\right) < W^P\left(n_I^N(\mu), \mu\right) < W^P\left(n_I^P(\mu), \mu\right) < W^P\left(n_*^P(\mu), \mu\right)$$

This proves that $\tilde{\mu}_0$ is unique and $W^P\left(n_I^P(\mu), \mu\right) \lesseqgtr W^N\left(n_I^N(\mu), \mu\right)$ as $\mu \lesseqgtr \tilde{\mu}_0$.

Proof of Proposition 5. At the second stage, the ISP chooses the regime independently of coverage, so that the regime is N if $\mu \leq \mu_R$ and P if $\mu > \mu_R$. If $\mu \leq \mu_R$ or $\mu > \mu_B$, the regime chosen by the ISP is also the regime preferred by the regulator, so that the regulator can impose the welfare-maximizing coverage. If $\mu \in (\mu_R, \mu_B]$, the ISP chooses P while N is the welfare-maximizing regime, so that the regulator chooses $n_*^P < n_*$.

Proof of Proposition 6. First, we prove that there exists a $\tilde{\mu}_1 < \mu_R$ such that $B^P(\tilde{\mu}_1) - R^N(\tilde{\mu}_1) = \Delta V(\tilde{\mu}_1) + aX^P(\tilde{\mu}_1) = 0$, and $n_*^P = n_I^N$. Indeed, consider a $\mu < \mu_R$ and $\Delta V(\mu) + aX^P(\mu) > \Delta V(\mu) + a\Delta X_0(\mu) = \Delta R(\mu)$. Now, since $\Delta V(\mu) < 0$ for $1 < \mu < \mu_R$, $\Delta V(1) + aX^P(1) = 0$, $\Delta V(\mu_R) + aX^P(\mu_R) = B^P(\mu_R) - R^N(\mu_R) > \Delta R(\mu_R) = 0$ and the fact that $X^P(\mu) = \frac{\mu^2 - 1}{2(\alpha+\beta)}\left(\alpha + \frac{\beta}{\mu}\right)$ is strictly increasing, as $X^{P\prime}(\mu) = \frac{\beta + 2\alpha\mu^3 + \beta\mu^2}{\mu^2} > 0$. This completes the proof.

Second, we turn to cases (a)-(e).

(a) if $\mu < \tilde{\mu}_1$, then $R^N(\mu) > B^P(\mu)$ so that $n_I^N > n_*^P$, which implies that $n_*^N > n_I^N > n_*^P > n_I^P$. As $n_I^N > n_I^P$, this also implies that N is chosen under UM and USO. Moreover, since $\mu < \mu_R$, we have $R^N(\mu) > R^P(\mu)$, so that

$$W^N\left(n_I^N, \mu\right) > W^P\left(n_I^P, \mu\right)$$

and N is chosen under TMR. Since $\mu < \mu_B$, we have $B^N(\mu) > B^P(\mu)$, so that

$$W^N\left(n_*^N, \mu\right) > W^P\left(n_*^P, \mu\right)$$

so that N is chosen under FR.

(b) if $\tilde{\mu}_1 \leq \mu < \mu_R$ then $B^P(\mu) \geq R^N(\mu)$ so that $n_*^P \geq n_I^N$. Since $\mu < \mu_B$, $n_*^N > n_*^P$ and since $\mu < \mu_R$, $n_I^N > n_I^P$, We thus have that

$$n_*^N > n_*^P \geq n_I^N > n_I^P$$

Since $\mu < \mu_B < \mu_B$, arguments in (a) apply to show that N is chosen under all regulatory frameworks.

(c) if $\mu_R \leq \mu < \tilde{\mu}_0$ then $W^N\left(n_I^N, \mu\right) > W^P\left(n_I^P, \mu\right)$, so that N is chosen under TMR. Since $\mu \geq \mu_R$, $n_I^P \geq n_I^N$ and $\Pi^P\left(n_I^P, \mu\right) \geq \Pi^N\left(n_I^N, \mu\right)$. Since $\mu < \mu_B$, $n_*^N > n_*^P$ and $W^N\left(n_*^N, \mu\right) > W^P\left(n_*^P, \mu\right)$. We thus have that

$$n_*^N > n_*^P > n_I^P \geq n_I^N$$

and that P is chosen under UM and USO, while N is chosen under FR.

(d) $\tilde{\mu}_0 \leq \mu < \mu_B$, then $W^P\left(n_I^P, \mu\right) > W^N\left(n_I^N, \mu\right)$, so that P is chosen under TMR. Since $\mu_R > \mu > \mu_B$, coverages are set as in (c). Arguments in (c) apply to show that P is chosen under UM and USO, while N is chosen under FR.

(e) if $\mu \geq \mu_B$ then $n_*^P \geq n_*^N$ and $W^P\left(n_*^P, \mu\right) \geq W^N\left(n_*^N, \mu\right)$ and P is chosen under FR. Since $\mu > \tilde{\mu}_0$, $n_*^P > n_I^N$ and since $\mu > \mu_R$, $n_I^P > n_I^N$ and

$$n_*^P \geq n_*^N > n_I^P > n_I^N$$

and P is also chosen under UM, USO and TMR.

Proof of Proposition 7. For $\mu \in (\tilde{\mu}_0, \mu_B)$, we have $n_U = n_*^P > n_T = n_I^P$ so this yields $W^P\left(n_*^P(\mu), \mu\right) > W^P\left(n_I^P(\mu), \mu\right)$, which proves the second part. If $\mu \in (\mu_R, \tilde{\mu}_0]$, from Proposition 6, $n_*^N > n_*^P > n_I^P \geq n_I^N$ and if $\mu = \tilde{\mu}_0$: $W^P\left(n_*^P(\tilde{\mu}_0), \tilde{\mu}_0\right) > W^P\left(n_I^P(\tilde{\mu}_0), \tilde{\mu}_0\right) = W^N\left(n_I^P(\tilde{\mu}_0), \tilde{\mu}_0\right) > W^N\left(n_I^N(\tilde{\mu}_0), \tilde{\mu}_0\right)$. Now we have

$$\begin{aligned}
& W^P\left(n_*^P(\mu), \mu\right) - W^N\left(n_I^N(\mu), \mu\right) \\
= \; & \Pi^P\left(n_*^P(\mu), \mu\right) - \Pi^N\left(n_I^N(\mu), \mu\right) + a\left(X_1^P(\mu) - X_1^N(\mu)\right) \\
< \; & \Pi^P\left(n_I^P(\mu), \mu\right) - \Pi^N\left(n_I^N(\mu), \mu\right) + a\left(X_1^P(\mu) - X_1^N(\mu)\right)
\end{aligned}$$

as $\Pi^P\left(n_I^P(\mu), \mu\right) > \Pi^P(n, \mu)$ for all n. So if $\mu = \mu_R$

$$W^P\left(n_*^P(\mu_R), \mu_R\right) - W^N\left(n_I^N(\mu_R), \mu_R\right) < a\left(X_1^P(\mu_R) - X_1^N(\mu_R)\right) < 0$$

since $X_1^P(\mu) - X_1^N(\mu) = -\frac{\beta}{2}\frac{(\mu-1)^2}{\mu(\alpha+\beta)} < 0$. By continuity, there exists a $\tilde{\mu}_u \in (\mu_R, \tilde{\mu}_2)$ such that $W^P\left(n_*^P(\tilde{\mu}_u), \tilde{\mu}_u\right) = W^N\left(n_I^N(\tilde{\mu}_u), \tilde{\mu}_u\right)$ and $W^P\left(n_*^P(\mu), \mu\right) \lesseqqgtr W^N\left(n_I^N(\mu), \mu\right)$ as $\mu \lesseqqgtr \tilde{\mu}_u$.

Proof of Lemma 2. *First* let us give a characterization of M_*^i. Putting first-best coverages $n_*^i(\mu) = \frac{B^i(\mu)}{c\mu}$ in the ISP participation constraint (14), implies that bandwidth levels are such that $R^i(\mu) - \frac{1}{2}B^i(\mu) \geq 0$. This can also be written as $R^i(\mu) - aZ^i(\mu) \geq 0$, where $Z^N(\mu) = X^N(\mu)$, $Z^P(\mu) = X_0^N(\mu) + X_1^P(\mu)$. Let $\Delta Z(\mu) \equiv X_1^P(\mu) - X_1^N(\mu) = -\frac{\beta}{2}\frac{(\mu-1)^2}{\mu(\alpha+\beta)} \leq 0$, which is decreasing and concave for all $\mu > 1$. Here $Z^N(\mu)$ is increasing and linear and $Z^P(\mu)$ is increasing and concave with μ. Then for $a = 0$, we have $R^i(\mu) - aZ^i(\mu) \geq 0$, where $Z^N(\mu) \geq 0$ for all $\mu \geq 1$. Moreover, $\lim_{\mu\to+\infty}\left(R^i(\mu) - aZ^i(\mu)\right) = \lim_{\mu\to+\infty}\left(V^i(\mu) - a\hat{Z}^i(\mu)\right)$ where $\hat{Z}^N(\mu) = X^N(\mu)$ and $\hat{Z}^P(\mu) = X^P(\mu) - 2X_0^N(\mu) = \frac{1}{2}(\mu-1)\frac{\alpha\mu^2+(\beta-3\alpha)\mu+\beta}{\mu(\alpha+\beta)} > 0$.[32] So

$$\begin{aligned}
\lim_{\mu\to+\infty}\left(V^N(\mu) - aX^N(\mu)\right) &= \lim_{\mu\to+\infty}\left(v(\mu-1)^{\alpha+\beta} - a(\mu-1)\right) = \lim_{\mu\to+\infty}(-a\mu) = -\infty \\
\lim_{\mu\to+\infty}\left(V^P(\mu) - a\hat{Z}^P(\mu)\right) &= \lim_{\mu\to+\infty}\left(\left(\frac{1}{2}\right)^{\alpha+\beta} v\left(\mu^2-1\right)^{\alpha+\beta}\mu^{-\beta} - a\frac{1}{2}(\mu-1)\frac{\alpha\mu^2+(\beta-3\alpha)\mu+\beta}{\mu(\alpha+\beta)}\right) \\
&= \lim_{\mu\to+\infty}\left(v\left(\frac{1}{2}\right)^{\alpha+\beta}\mu^{2\alpha} - a\frac{\alpha\mu^2}{2(\alpha+\beta)}\right) = \lim_{\mu\to+\infty}\left(-a\frac{\alpha\mu^2}{2(\alpha+\beta)}\right) = -\infty
\end{aligned}$$

So it exists a unique ad price $a_\pi^i : \pi^i(\mu) = 0$. More precisely

$$\begin{aligned}
a_\pi^N &= \frac{V^N(\mu)}{X^N(\mu)} = v(\mu-1)^{\alpha+\beta-1} > 0 \\
a_\pi^P &= \frac{V^P(\mu)}{\hat{Z}^P(\mu)} = v(\alpha+\beta)\left(\frac{1}{2}\right)^{\alpha+\beta-1}\frac{(\mu-1)^{\alpha+\beta-1}(\mu+1)^{\alpha+\beta}\mu^{-(1+\beta)}}{\alpha\mu^2+(\beta-3\alpha)\mu+\beta} > 0
\end{aligned}$$

[32] Indeed, the quadratic polynom $\alpha\mu^2 + (\beta-3\alpha)\mu + \beta$ has no real roots if $\beta < 9\alpha$, and only negative roots if $\beta \geq 9\alpha$.

When $\alpha+\beta<1$, these price thresholds are decreasing functions of μ with $\lim_{\mu\rightarrow+\infty} a_{\pi}^{i}=0$, so that for all a, it exists a unique $\bar{\mu}^{i}:a_{\pi}^{i}=a$. Then $\bar{\mu}^{i}$ becomes the inverse function of a_{π}^{i} and it is decreasing in a. As a result the ISP participation constraint (14) is satisfied for $\mu\leq\bar{\mu}^{i}$ and $M_{*}^{i}=[1,\bar{\mu}^{i}]$. [Utilité du second point ???]*Second*, we see that

$$\begin{aligned}
\Delta a_{\pi} &\equiv a_{\pi}^{P}-a_{\pi}^{N}\\
&= v\left(\mu-1\right)^{\alpha+\beta-1}\left[\left(\alpha+\beta\right)\left(\frac{1}{2}\right)^{\alpha+\beta-1}\frac{\left(\mu+1\right)^{\alpha+\beta}\mu^{-(1+\beta)}}{\alpha\mu^{2}+\left(\beta-3\alpha\right)\mu+\beta}-1\right]
\end{aligned}$$

so $\lim_{\mu\rightarrow1^{+}}\Delta a_{\pi}=\lim_{\mu\rightarrow1^{+}}2\frac{\alpha v}{\beta-\alpha}\left(\mu-1\right)^{\alpha+\beta-1}=0^{+}$ and

$$\begin{aligned}
\lim_{\mu\rightarrow+\infty}\Delta a_{\pi} &= \lim_{\mu\rightarrow+\infty}v\left(\mu-1\right)^{\alpha+\beta-1}\left[\left(\frac{\alpha+\beta}{\alpha}\right)\left(\frac{1}{2}\right)^{\alpha+\beta-1}\mu^{-(3+\alpha)}-1\right]\\
&= \lim_{\mu\rightarrow+\infty}-v\left(\mu-1\right)^{\alpha+\beta-1}=0^{-}
\end{aligned}$$

So it exists (at least) one $\mu_{a}:\Delta a_{\pi}=0$ such that $a_{\pi}^{P}\geq a_{\pi}^{N}$ if $\mu\leq\mu_{a}$ and conversely.

Proof of Proposition 8. Proof of Proposition 6 applies with the restrictions that $\tilde{\mu}_{1}=\mu_{R}$ and $n_{*}^{i}=n_{\pi}^{i}$.

Proof of Proposition 9. Proof of Proposition 6 applies by replacing $\tilde{\mu}_{1}$ by $\hat{\mu}$ and imposing the restriction that $\mu_{B}=\tilde{\mu}_{0}=\mu_{R}$.

Proof of Proposition 10. First, as we have $B_{1}^{i}\left(\mu\right)<B^{i}(\mu)$ for $i=N,P$, maximum USO coverages are $n_{f}^{i}=\frac{B_{1}^{i}(\mu)}{c\mu}<n_{*}^{i}=\frac{B^{i}(\mu)}{c\mu}$. Second, as $B_{1}^{i}\left(\mu\right)-R^{i}(\mu)=aX_{1}^{i}\left(\mu\right)>0$ then $n_{f}^{i}>n_{I}^{i}$ for all $\mu\geq1$. Third, since $\Delta B_{1}=\Delta B$, $n_{f}^{P}\lesseqqgtr n_{f}^{N}$ as $\mu\lesseqqgtr\mu_{B}$. Fourth, as $B_{1}^{P}\left(\mu\right)-R^{N}(\mu)<B^{P}\left(\mu\right)-R^{N}(\mu)$ and $B^{P}(\tilde{\mu}_{f})>R^{N}(\tilde{\mu}_{f})$, similar arguments as in the proof of Lemma **??** applies, so that there exists a $\tilde{\mu}_{f}\in(\tilde{\mu}_{1},\mu_{R})$ such that $B_{1}^{P}\left(\tilde{\mu}_{f}\right)=R^{N}(\tilde{\mu}_{f})$. Then, for $\mu\leq\tilde{\mu}_{f}$, $n_{f}^{P}\leq n_{I}^{N}$. Furthermore, as $W_{1}^{P}\left(n_{I}^{P},\mu\right)-W_{1}^{N}\left(n_{I}^{N},\mu\right)=n_{I}^{P}B_{1}^{P}(\mu)-C\left(n_{I}^{P},\mu\right)-n_{I}^{N}B_{1}^{N}(\mu)+C\left(n_{I}^{N},\mu\right)$, we follow same steps as in the proof of Proposition 4 to show that, by continuity, there exists a $\hat{\mu}_{f}\in(\mu_{R},\mu_{B})$ such that $W_{1}^{P}\left(n_{I}^{P},\hat{\mu}_{f}\right)=W_{1}^{N}\left(n_{I}^{N},\hat{\mu}_{f}\right)$. Moreover since $W_{1}^{P}\left(n_{I}^{P},\mu\right)-W_{1}^{N}\left(n_{I}^{N},\mu\right)=W^{P}\left(n_{I}^{P},\mu\right)-W^{N}\left(n_{I}^{N},\mu\right)-aX_{0}^{N}(n_{I}^{P}-n_{I}^{N})$, then we have $W_{1}^{P}\left(n_{I}^{P},\tilde{\mu}_{0}\right)-W_{1}^{N}\left(n_{I}^{N},\tilde{\mu}_{0}\right)=-aX_{0}^{N}(n_{I}^{P}-n_{I}^{N})<0$, because $n_{I}^{P}>n_{I}^{N}$ as $\tilde{\mu}_{0}>\mu_{R}$. This implies that $\hat{\mu}_{f}>\tilde{\mu}_{0}$. Finally, other arguments are unchanged from Proposition 6 and are omitted in this proof. Then Proposition 6 applies for n_{f}^{i} and n_{I}^{i} by substituting n_{f}^{i} to n_{*}^{i}, $\tilde{\mu}_{f}$ to $\tilde{\mu}_{1}$ and $\hat{\mu}_{f}$ to $\tilde{\mu}_{0}$.

Proof of Lemma 3. From Proposition 1, $\mu_R(1) < \mu_B$. Since $\Delta R = \Delta V$ at $\eta = 0$, we have $\mu_R(0) = \mu_V$, which is greater than μ_B from Lemma 1 and Proposition 1. By continuity there exists a $\hat{\eta} \in (0,1)$ such that $\mu_R(\eta) \lesseqqgtr \mu_B$ as $\eta \gtreqqless \hat{\eta}$.

Proof of Proposition 11. We assume $\alpha < 0.5$. Note that

$$\Delta B - \Delta R \lesseqqgtr 0 \text{ as } \eta \gtreqqless \hat{\eta}$$

If $\eta > \hat{\eta}$, the proof of Proposition 1 applies once we substitute $\mu_R(\eta)$ to μ_R in it.

Consider now the case where $0 < \eta \leq \hat{\eta}$. Since $\Delta V \lesseqqgtr 0$ as $\mu \lesseqqgtr \mu_V$, $a\eta\Delta X_0(1) = 0$, $a \lim_{\mu\to\infty} \eta\Delta X_0 \to \infty$ and $\frac{\partial(\Delta X_0)}{\partial\mu} > 0$, $\forall\mu$, there exists a $\mu_R(\eta) < \mu_V$ such that $\Delta R = \Delta V + \eta a \Delta X_0 \lesseqqgtr 0$ as $\mu \lesseqqgtr \mu_R(\eta)$.

Similarly, since $\Delta R \lesseqqgtr 0$ as $\mu \lesseqqgtr \mu_R(\eta) > 0$, $a(\Delta X(1)) = 0$, $\lim_{\mu\to\infty} a(\Delta X(\mu)) \to \infty$, $a(\Delta X(\mu))$ is convex and $\Delta B \geq 0 = \Delta R$ at $\mu = \mu_R(\eta)$, there exists a $\mu_B \leq \mu_R(\eta)$ such that $\Delta B \lesseqqgtr 0$ as $\mu \lesseqqgtr \mu_B$.

Proof of Proposition 12. If $\eta > \hat{\eta}$, the proof of Proposition 6 applies once we substitute $\mu_R(\eta)$ to μ_R in it. Consider now the case of $\eta < \hat{\eta}$.

(a) The case $\mu < \tilde{\mu}_1(\eta)$ is identical to case (a) of Proposition 6 because μ is less than all bandwidth threshold.

(b) The case $\tilde{\mu}_1(\eta) \leq \mu < \mu_B < \mu_R(\eta)$ is identical to case (b) of Proposition 6 as μ is lower than both μ_B and μ_R.

(c) If $\mu_B \leq \mu < \tilde{\mu}_0$, then, because $\mu \geq \mu_B$, $n_*^P > n_*^N$ and $W^P\left(n_*^P, \mu\right) \geq W^N\left(n_*^N, \mu\right)$ so that P meets is chosen under FR. But, because $\mu < \tilde{\mu}_0$, $W^N\left(n_I^N, \mu\right) > W^P\left(n_I^P, \mu\right)$ and N is chosen under TMR. Since $\mu < \mu_R(\eta)$, $n_I^N > n_I^P$ and N is chosen under UM and USO. We thus have $n_*^P > n_*^N > n_I^N > n_I^P$.

(d) If $\tilde{\mu}_0 \leq \mu < \mu_R(\eta)$, $W^P\left(n_I^P, \mu\right) \geq W^N\left(n_I^N, \mu\right)$ and P is chosen under TMR. Arguments in (c) apply for FR,USO and UM.

(e) The case $\mu \geq \mu_R(\eta)$, is identical to case (e) of Proposition 6 because because μ is greater than all bandwidth threshold.

Proof of Proposition 13. From (8) and (21), we obtain:

$$
\begin{aligned}
m'_n &= \frac{1}{2}\left(\frac{aX_f^i}{F}\right)^{\frac{1}{2}} n^{-\frac{1}{2}} > 0 \\
m'_\mu &= \frac{1}{2}\left(\frac{an}{F}\right)^{\frac{1}{2}} \left(X_f^i\right)^{-\frac{1}{2}} \left(X_f^i\right)' > 0 \\
m''_{\mu n} &= \frac{1}{4}\left(\frac{an}{F}\right)^{-\frac{1}{2}} \left(X_f^i\right)^{-\frac{1}{2}} \left(X_f^i\right)' > 0
\end{aligned}
$$

Moreover, since

$$
\Delta m\left(n,\mu\right) = \left(\frac{an}{F}\right)^{\frac{1}{2}} \left(\sqrt{X_f^P\left(\mu\right)} - \sqrt{X_f^N\left(\mu\right)}\right) < 0
$$

we obtain:

$$
\begin{aligned}
\Delta m\left(n,\mu\right) &= \left(\frac{na}{F}\right)^{\frac{1}{2}} \left(\sqrt{X_f^P\left(\mu\right)} - \sqrt{X_f^N\left(\mu\right)}\right) = \left(\frac{na}{F}\right)^{\frac{1}{2}} \left(\sqrt{\frac{1}{2}\frac{\beta}{\alpha+\beta}\frac{\left(\mu^2-1\right)}{\mu}} - \sqrt{\frac{\beta}{\alpha+\beta}\left(\mu-1\right)}\right) < 0 \\
\frac{\partial\left(\Delta m\right)}{\partial n} &= \frac{1}{2}n^{-1}\Delta m < 0 \\
\frac{\partial^2\left(\Delta m\right)}{\partial n^2} &= -\frac{1}{4}n^{-2}\Delta m > 0
\end{aligned}
$$

List of Figures

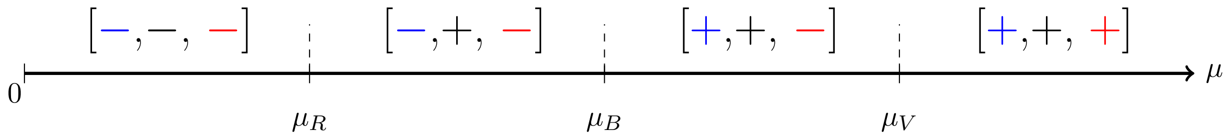

Figure 1: Social welfare ΔB, ISP profit ΔR and consumer surplus ΔV differences

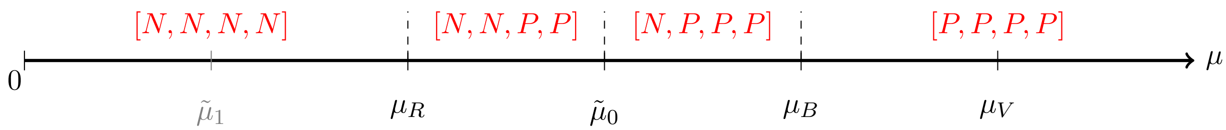

Figure 2: Net Neutrality or Prioritization : criteria [FR, TMR, USO, UM]

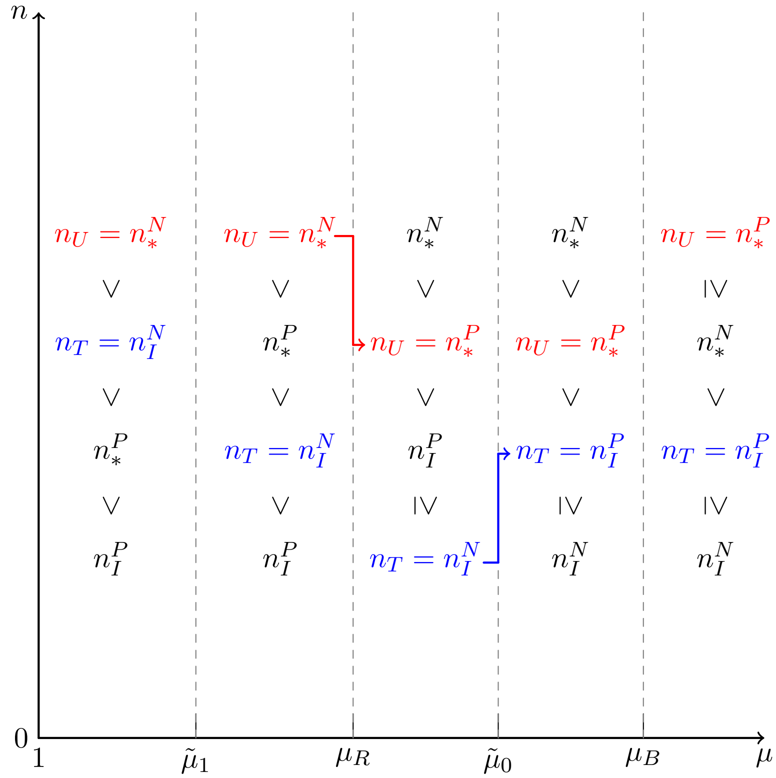

Figure 3: Proposition 6. n_T: in blue n_U: in red

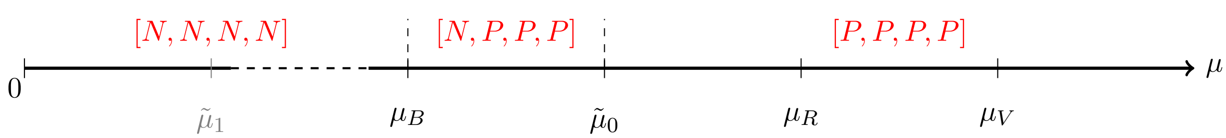

Figure 4: Criteria with ISP low bargaining power