Review of Industrial Organization
https://doi.org/10.1007/s11151-024-09945-1

# Net Neutrality and Universal Service Obligations: It's All About Bandwidth

**Axel Gautier**[1,2] · **Jean-Christophe Poudou**[3] · **Michel Roland**[4]



**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 subject 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: She can mandate the broadband market coverage. We show that the greater is the network bandwidth, the more likely is the repeal of NN to increase 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 Classification** D21 · K23 · L12 · L51 · L96

✉ Axel Gautier
agautier@uliege.be

Jean-Christophe Poudou
jean-christophe.poudou@umontpellier.fr

Michel Roland
Michel.Roland@ecn.ulaval.ca

1 LCII, HEC Liege, University of Liege, Liege, Belgium

2 CESifo, Munich, Germany

3 MRE, University of Montpellier, Montpellier, France

4 CREATE, Département d'économique, Université Laval, Québec, Canada

Published online: 06 March 2024

## 1 Introduction

Most countries impose two types of regulation on Internet Service Providers (ISPs): Net Neutrality (NN) and Universal Service Obligations (USOs). The former prohibits ISPs from speeding up, slowing down or blocking internet traffic based on its source, ownership or destination (Krämer et al., 2013). It aims to promote investment, innovation and competition among content providers (CPs) and (more generally) to ensure free speech (Katz, 2017). The latter forces ISPs to cover a given percentage of the territory with a minimum broadband standard in terms of download and upload speeds. Its goal is to avoid a "digital divide" among citizens of different regions (McMenemy, 2022).

Because of the growth of data-intensive content on the internet, ISPs argue that it is nowadays counterproductive to treat in the same way the CPs that require high speed of transmission and that do not tolerate delays—such as streaming—as those that are far less demanding on those counts: such as emails (Peitz and Schuett, 2016). As a result, there are debates within regulatory agencies and among academics with respect to the ongoing relevance of NN. For instance, in the US, the Federal Communications Commission (FCC) implemented NN in 2015, repealed it in 2018, and proposed to reinstate it in October 2023 (Congressional Research Service, 2023). In contrast, there is a clear tendency to strengthen USOs almost everywhere in the world, including in the US (Garci-Calvo, 2012).

A striking feature of the debates on internet regulation is that they treat NN and USO policies independently. Under NN, one can analyze internet USOs with the same methodological approach as has been applied to other industries that have already been 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.

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 since both types of regulation directly affect internet service pricing and the incentives to invest in broadband networks.

In this paper, we analyze both NN and USOs in a single model. Our objective is to understand the relations between the two policies and how they interact. In particular, the choice of NN or P can be viewed as a choice between a one-sided USO versus a two-sided USO. This interplay is the key element in our analysis.

We consider a two-sided market where a monopolistic ISP can install a broadband network of a given bandwidth at different locations in a country. A regulator can determine whether to submit the ISP to NN or to allow it to supply paid priority services to CPs with the objective 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 adversely affect consumers if they have strong preferences for the non-prioritized service.

The overall effect depends on the network bandwidth: The greater is the bandwidth—the greater is the network data transfer rate—the lesser is the impediment of prioritization 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 affects welfare but not the ISP profit, there is a bandwidth range over which the repeal of NN increases the ISP profit while it decreases welfare. A regulation that mandates NN can then improve welfare.

With respect to the USO literature,[1] we use the standard framework in which a regulator determines the extent of a total market that the ISP network must cover; 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. A recurrent theme is to evaluate the welfare effect of a uniform pricing constraint, which is a ban on third-degree price discrimination.[2] To our knowledge, only one-sided markets have been analyzed. We instead consider a pricing constraint in a two-sided market: NN, which is a ban on third-degree price discrimination on the CP side of the market.

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 (Bourreau et al., 2017; Choi and Kim, 2010; Reggiani and Valletti, 2016). We thus refer to a case where bandwidth is determined primarily by the current state of technology. 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 an 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 25/3 Mbps in 2015,[3] and this is probably called for a revision soon.[4] Broadband definitions also vary across countries.[5]

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.,

[1] Early contributions are Anton et al. (2002) and Valletti et al. (2002).

[2] Valletti et al. (2002) and Gautier and Wauthy (2010) study the impact of uniform pricing on entry and the extent of competition; Poudou and Roland (2014) provides an efficiency justification for imposing uniform pricing constraints, while Poudou and Roland (2017) study the consequences on inequality among consumers.

[3] See BroadbandNow (2021).

[4] On November 1, 2023, the FCC launched a notice of inquiry to increase the national fixed broadband speed benchmark to 100/20 Mbps (FCC, 2023)

[5] For instance, the EU defines a 30 Mbps download speed as fast broadband and a 100 Mbps as ultrafast broadband (Bourreau et al., 2017). Coverage targets are given in both terms. Canada sets broadband coverage targets in terms of 50/10 Mbps download/upload speeds (CRTC, 2021).

2021). 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.

In the next section, we present the model of the two-sided internet market that we analyze, and we specify the way that NN and prioritization are defined and implemented. In Sect. 3, we perform the comparative statics between NN and prioritization, and we present the benchmark cases of welfare and profit-maximizing coverages. Section 4 provides the core results with respect to the choice between NN, and prioritization as well as results with respect to the determination of market coverage as a function of bandwidth. For ease of presentation, these results are obtained under simplifying assumptions, but we present extensions in Sect. 5. The conclusion in Sect. 6 sums up the main results of our model.

## 2 Basic Model

We consider a country that is composed of a continuum of locations $n \in [0, \infty[$ that are ranked in increasing order of network deployment cost. A monopolistic ISP connects consumers to content providers in covered locations. A regulator oversees the ISP with the objective of maximizing social welfare within the confines of the available regulatory tools.

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.[6]

### 2.1 Content Providers (CPs)

There are two types of content providers denoted by $j = 0, 1$. We consider that CP 0 is a large content provider (streaming platform, social network, etc.), while CP 1 is a fringe of small content (newspapers, info, blogs, etc.).

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; we assume that $a$ is the same for the two CPs. Operating costs are normalized to zero. We denote by $X_j$ the total traffic per location (in MB) of CP $j$. A CP's revenue is $\Pi_j = anX_j$. We let $X \equiv X_0 + X_1$ be the total traffic per location.

### 2.2 Internet Service Provider (ISP)

The ISP operates a network of bandwidth $\mu > 1$ to link CPs to consumers in $n$ locations. The bandwidth size $\mu$ is the same in all locations. The maximum traffic that

[6] All of the variables that are used in our model are reported in the nomenclature of variables in Table 2.

the ISP ne twork can handle depends on bandwidth and on the traffic management that it practices in the event 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]

We consider two traffic management practices: Net neutrality (N), where traffic is managed under a best-efforts service and the average content delivery delay is identical for all content; and Prioritization (P), where prioritized content is given precedence in the event of congestion. Under P, there is a fast and a slow lane for routing content. Since bandwidth is a scarce resource, a heavy-bandwidth-using CP slows down the other CP and thereby imposes a negative externality.

Content providers can pay the ISP to have their content carried on the fast lane, while the slow lane remains free.[8] As we consider that CP 1 is a fringe of small producers while CP 0 is a large platform, the ISP will contract for prioritization with CP 0. The cost of establishing a network that covers markets $[0, n]$ with a bandwidth $\mu$ is:

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

so that the marginal cost of coverage is increasing with bandwidth.

### 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:

$$U\left(X_0, X_1\right) = X_0^{\alpha} X_1^{\beta}, \tag{2}$$

where $\alpha + \beta < 1$.

Let $\left(X_0^i(\mu), X_1^i(\mu)\right)$, $i = N, P$, be the content consumption under $i$. We define the indirect utility function in regime $i$ as $V^i(\mu) \equiv U^i\left(X_0^i(\mu), X_1^i(\mu)\right)$.

### 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 $p^i = V^i(\mu)$.

[7] Choi et al. (2015, p. 452). The M/M/1 queue system is also used in Choi and Kim (2010), Krämer and Wiewiorra (2012), Bourreau et al. (2015), Reggiani and Valletti (2016) and Choi et al. (2018).

[8] The reason is that a CP should not need a contract with the ISP to be delivered to consumers. Usually content is stored in data warehouses that are outside of the ISP's network, and the ISP cannot refuse to deliver the content that is requested by the consumers. Legal content cannot be blocked by the ISP. In this context, the ISP cannot charge content providers for being in the slow lane: The basic service should necessarily be free.

**Table 1** Regulatory frameworks

| Regulation | Party that chooses coverage | Party that chooses traffic management |
|---|---|---|
| TMR | ISP | Regulator |
| USO | Regulator | ISP |
| FR | Regulator | Regulator |
| UM | ISP | ISP |

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 the prioritized content provider (CP 0) has 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. Under prioritization, the ISP per location revenue is then:

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

A benevolent regulator monitors this two-sided market. Depending on the regulatory framework considered, she 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. In all cases, we suppose that traffic management and coverage are decided simultaneously. Table 1 summarizes the responsibilities of the ISP and the regulator under each regulatory framework.

We gauge the performance of these regulatory frameworks in terms of coverage and social welfare against the benchmark cases of an 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, where the welfare-maximizing regime and coverage are considered notwithstanding any market or institutional constraint.

## 2.5 Traffic Management and Content Consumption

### 2.5.1 Net Neutrality (N)

With net neutrality, the average content delivery delay given by the M/M/1 queue model is:

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

where $\lambda$ is the frequency (in $s^{-1}$) of data transmission to the network.[9]

This equation determines the amount of traffic that could be delivered at an average waiting time $\bar{\omega}$, or alternatively at an average speed of transmission $\frac{1}{\bar{\omega}}$:

$$X = \left(\mu - \frac{1}{\bar{\omega}}\right)\frac{1}{\lambda}.$$

The ISP announces an average transmission speed $\frac{1}{\bar{\omega}}$ in MB/s. Normalizing the transmission speed to $\frac{1}{\bar{\omega}} = 1$ and taking $\lambda = 1$, the network has thus a capacity constraint that is given by:

$$X_0 + X_1 = \mu - 1. \tag{3}$$

Under N, the consumers' problem is:

$$\max_{X_0, X_1} U(X_0, X_1) \text{ subject to (3).} \tag{4}$$

The traffic management limits content consumption: Any increase of consumption above the level determined by the constraint (3 ) cannot be delivered by the ISP at the announced delivery speed and cannot be served by the network. In our formulation, constraint (3) can be interpreted as a time budget constraint that is faced by the consumer.

The solution to problem (4) gives the content consumption under N that we express as a function of the bandwidth capacity:

$$X_0^N(\mu) = \frac{\alpha}{\alpha + \beta}(\mu - 1),\ X_1^N(\mu) = \frac{\beta}{\alpha + \beta}(\mu - 1). \tag{5}$$

### 2.5.2 Prioritization (P)

The ISP can alternatively route traffic with a prioritization system under which the bandwidth is split into two lanes: a fast lane for prioritized traffic; and a slow lane. More precisely, a fraction $\rho < 1$ of the traffic $X = \mu - 1$ that is observed under net

[9] A more plausible alternative, but at the cost of tedious complications, would be that the average waiting time is dependent of the number of locations: $\bar{\omega} = 1/(\mu - \lambda n X)$. In this case, an increase in $n$ increases congestion; it means that the budget constraint is stronger when $n$ expands, so that the (derived below) utility functions $V^i$ decrease with $n$. As a consequence, optimal coverages for both the regulator and the ISP are lower.

neutrality is given precedence in the event of congestion: The ISP specifies a maximum amount of traffic that can be handled in the prioritized service. We assume that $\rho$ is greater than the share $\frac{\alpha}{\alpha+\beta}$ of $X_0^N$ in total traffic under net neutrality.

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:

$$\omega_0 = \frac{1}{\mu - \rho(\mu - 1)} = \frac{1}{(1-\rho)\mu + \rho}; \text{ and} \tag{6}$$

$$\omega_1 = \mu\omega_0. \tag{7}$$

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 $\omega_0$ and $\omega_1$ weighted by the consumption shares in transmission equal 1:[10]

$$\omega_0 \cdot \frac{X_0}{\mu - 1} + \omega_1 \cdot \frac{X_1}{\mu - 1} = \bar{\omega} = 1. \tag{8}$$

Using (6) and (7) and multiplying both sides by $\frac{\mu-1}{\omega_0}$, we can write this as:

$$X_0 + \mu X_1 = (\mu - 1)((1-\rho)\mu + \rho). \tag{9}$$

Under P, the consumers' problem is:

$$\max_{X_0, X_1} U(X_0, X_1) \text{ subject to (9)}, \tag{10}$$

and consumption levels are given by:

$$X_0^P(\mu) = \frac{\alpha}{\alpha+\beta}(\mu - 1)((1-\rho)\mu + \rho),\ X_1^P(\mu) = \frac{\beta}{\alpha+\beta}\frac{(\mu-1)}{\mu}((1-\rho)\mu + \rho). \tag{11}$$

### 2.5.3 Bandwidth and Price Effect

From the demand functions, we can interpret a change from neutrality to prioritization as a simultaneous increase of "income" from $\mu - 1$ to $(\mu - 1)((1-\rho)\mu + \rho)$ and of the non-prioritized content "price" from 1 to $\mu$. We can accordingly decompose the effect of a change in $\mu$ into a bandwidth effect and a price effect. The bandwidth effect increases in $\mu$ and decreases in $\rho$; the price effect increases in $\mu$. Figure 1 illustrates these two effects for a case where the consumer is indifferent between net neutrality and prioritization.

If we compare the "time budget constraint" under N and P (Eqs. 3 and 9), prioritization implies a budget pivot around $(\rho(\mu - 1), (1-\rho)(\mu - 1))$. By a revealed

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

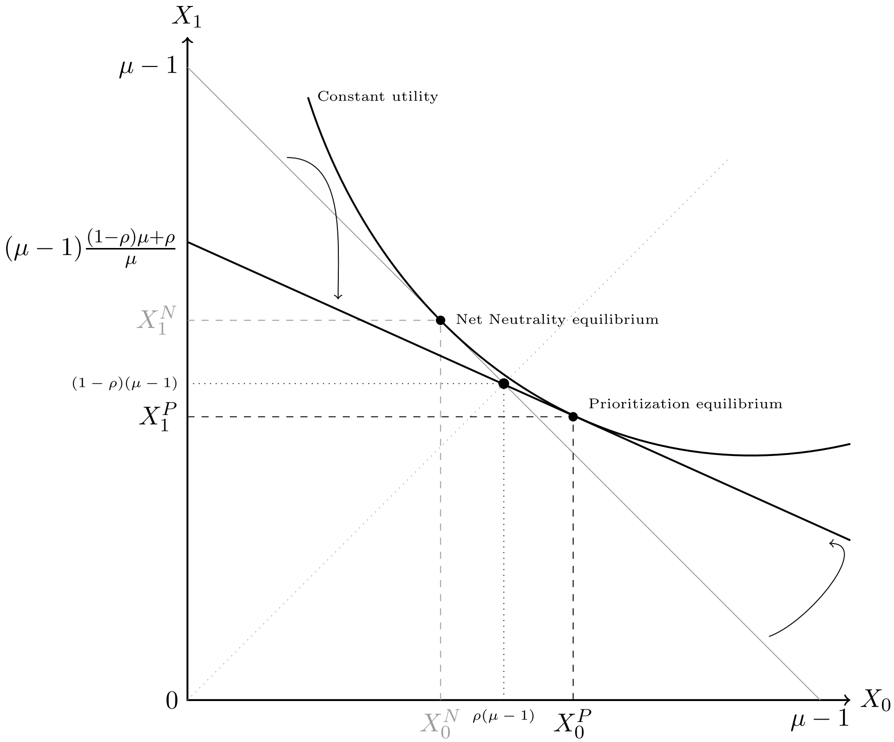


**Fig. 1** Bandwidth and the price effects of $\mu$

preference argument, consumers always prefer P to N if their consumption under N is an available choice under P. This is the case if either $\alpha > \beta$ or if $\rho = \frac{\alpha}{\alpha+\beta}$. In those circumstances, consumers and the ISP always prefer P to N and regulating traffic management is redundant. Therefore, the problem that we consider boils down to a classical USO problem where regulating market coverage is the only concern.

Prioritization involves a trade-off between a slower non-prioritized content and greater capacity only when consumers initially place more weight on the non-prioritized content. For this reason, hereafter we assume that $\alpha < \beta$ and $\rho > \frac{\alpha}{\alpha+\beta}$.

### 2.5.4 Indirect Utility Function, Revenue and Welfare

From (5) and (11), we obtain the following indirect utility functions.

$$V^N(\mu) = v(\mu-1)^{\alpha+\beta} \text{ ; and} \tag{12}$$

$$\begin{aligned} V^P(\mu) &= v(\mu-1)^{\alpha+\beta}\mu^{-\beta}((1-\rho)\mu+\rho)^{\alpha+\beta} \\ &= \frac{((1-\rho)\mu+\rho)^{\alpha+\beta}}{\mu^\beta} V^N(\mu) \end{aligned} \tag{13}$$

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

Based on these indirect utility functions, the ISP per location revenue and welfare functions have the following characteristics.

**Lemma 1** *For both traffic management practices $i = N, P$, for all bandwidth levels $\mu \geq 1$ and the unit benefit $a \geq 0$, we have*:

(a) $B^i(\mu) > R^i(\mu)$;
(b) $\frac{\partial B^i(\mu)}{\partial \mu} > \frac{\partial R^i(\mu)}{\partial \mu}$; *and*
(c) $\frac{\partial^2 B^i}{\partial a \partial \mu} > \frac{\partial^2 R^i}{\partial a \partial \mu}$.

This Lemma shows that for any level of bandwidth and whatever is the traffic management practice, total welfare is higher than the ISP revenue in each location, which is not surprising. This is also the case for their marginal values and for the marginal effect of a marginal increase in the unit traffic benefit.

## 3 Preliminary Results

In this section, we develop the fundamental results of the model that will underlie the analysis of the ISP's and regulator's choices.

### 3.1 Net Neutrality Versus Prioritization: Comparative Statics

As a first step, we compare market outcomes that occur under neutrality and prioritization for given $\mu$, $\rho$, 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 $H^i(\mu)$, $i = N, P$, we let $\Delta H \equiv H^P(\mu) - H^N(\mu)$.

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

**Lemma 2** *If $\alpha < \beta$, then*

(a) $\Delta X \geq 0$ *if and only if* $\mu \geq \mu_X \equiv \frac{\beta}{\alpha}\frac{1-\rho}{\rho}$;
(b) *There exists a* $\mu_V > \mu_X$ *such that* $V^P(\mu) \geq V^N(\mu)$ *if and only if* $\mu \geq \mu_V$; *and*

(c) *$\mu_X$ and $\mu_V$ are increasing in $\rho$.*

As in Economides and Hermalin (2012), the switch from N to P is similar to an increase of bandwidth and this potentially allows for more consumption i.e. it is similar to an increase in total income in a conventional consumer-choice model. However, P also changes the transmission speed, and this leads to the price effect. It makes one content relatively more attractive and the other relatively less attractive: the fast lane is slowing down the slow lane.[11] If $\mu \in [\mu_X, \mu_V]$, the change from neutrality to prioritization involves both an increase in total consumption and a decrease in utility because the income effect is not enough to compensate for the price effect.

In our analysis, we consider that the ISP can collect all of the additional revenue from the CP. If, instead, CP 0 has some bargaining power and the ISP can collect only a fraction ($\eta$) of the additional revenue, then $\Delta R = \Delta V + a\eta\Delta X_0 > 0$. A lower $\eta$ reduces the parameter space for which the ISP prefers P to N: $\mu_R$ decreases in $\eta$. If $\eta$ is low enough and the ISP can hardly monetize prioritization, we could have that $\Delta B \geq \Delta R$, which implies that $\mu_R > \mu_B$. Still there would be a parameter space where the ISP and the regulator are in conflict over the preferred traffic management practice.

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 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 for the lower revenue from the consumers. The ISP prefers prioritization when $\Delta R = \Delta V + a\Delta X_0 > 0$. This threshold from which prioritization increases the ISP's revenue is less than the threshold that increases welfare: Since $\Delta X_1 < 0$ and $\Delta B = \Delta R + a\Delta X_1 < \Delta R$, it takes a larger bandwidth to make prioritization improve welfare than the bandwidth that improves the ISP's revenue.

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

**Proposition 1** *There exist a $\mu_R$ 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. $\Delta R > \Delta B > 0, \forall \mu > \mu_B$; *and*
4. *$\mu_R$ and $\mu_B$ are increasing in $\rho$.*

[11] Economides and Hermalin (2012) do not consider this price effect. As a result, in their model, utility increases with total consumption.

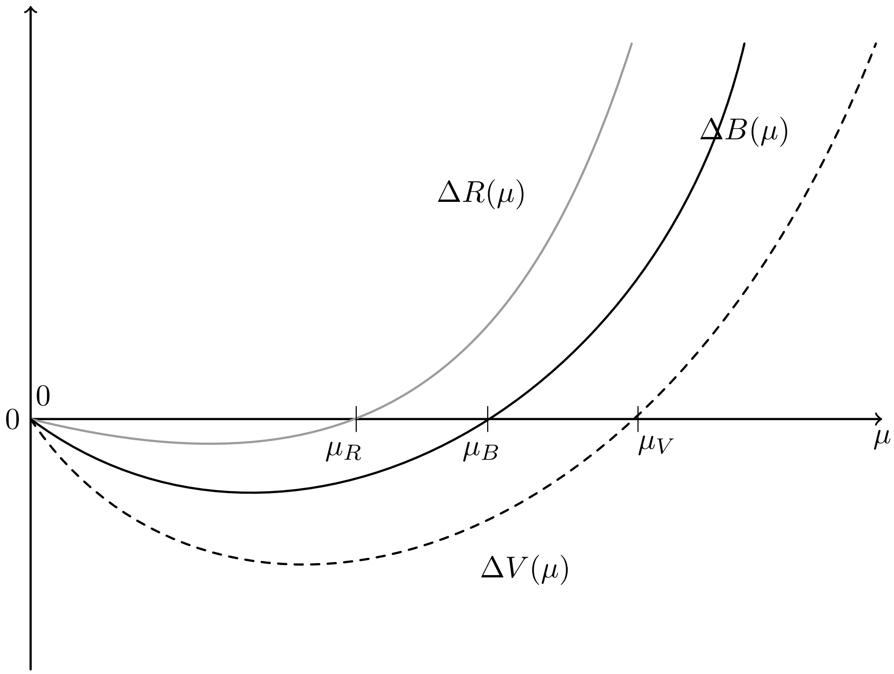


**Fig. 2** Social welfare (black), ISP profit (gray) and consumer surplus (dashed) differences

Proposition 1 implies that, for a given $\mu$ 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 prioritization regime. This is illustrated in Fig. 2.[12]

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.

According to both the previous Lemma and the previous Proposition, all of the important rankings that are needed for our analysis are not affected by the level of the splitting parameter $\rho$; accordingly, in order to simplify the presentation, we arbitrarily take a value of the parameter: $\rho = \frac{1}{2}$.

[12] The shape of the curves is given as an illustration.

### 3.2 Benchmark Coverages

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 (UM) Benchmark**

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:

$$R^i(\mu) - C_n'(n,\mu) = 0. \tag{14}$$

Denoting the solution by $n_I^i$, we obtain:

$$n_I^i(\mu) = \frac{R^i(\mu)}{c\mu}. \tag{15}$$

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:

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

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

As a result of Proposition 2, the profit-maximizing coverage is

$$n_I \equiv \max\left(n_I^N(\mu), n_I^P(\mu)\right).$$

and the profit-maximizing regime is $\arg\max_i n_I^i(\mu)$.[13]

**First-Best (FB) Benchmark**

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,

$$B^i(\mu) = C_n'(n,\mu).$$

[13] The profit maximizing coverage under P depends on the ability of the ISP to collect revenue from the prioritized CP. A lower $\eta$ reduces the ISP's revenue under P and therefore its coverage.

Denoting the solution by $n^i_*$, we obtain:

$$n^i_*(\mu) = \frac{B^i(\mu)}{c\mu}. \tag{16}$$

As $B^N(\mu) - R^N(\mu) = aX^N(\mu) > 0$ and $B^P(\mu) - R^P(\mu) = a\big(X^P_1(\mu) + X^N_0(\mu)\big) > 0$, social benefit is greater than ISP revenue under both management practices. As a result, the unregulated coverage is lower than the FB coverage for a given $\mu$. 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* $\mu \lesseqgtr \mu_B$,

$$\begin{aligned} n^P_*(\mu) &\lesseqgtr n^N_*(\mu); \quad \text{and} \\ W^P\big(n^P_*(\mu), \mu\big) &\lesseqgtr W^N\big(n^N_*(\mu), \mu\big). \end{aligned}$$

The FB coverage is thus

$$n_* \equiv \max\left(n^N_*(\mu), n^P_*(\mu)\right),$$

and the FB regime is $\arg\max_i n^i_*(\mu)$. Of course, the result that the FB coverage is greater than the profit-maximizing coverage is the basic feature of any model that involves universal service. But in our model, the regulator can act contrary not only to the ISP's preferred coverage but also to the ISP's 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

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: UM, TMR, USO, and FR. 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; 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 prioritization over net neutrality is lower than $\mu_B$. Moreover, this threshold is also greater than $\mu_R$ 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* $W^P(n_I^P(\tilde{\mu}_0), \tilde{\mu}_0) \lesseqqgtr W^N(n_I^N(\tilde{\mu}_0), \tilde{\mu}_0)$ *as* $\mu \lesseqqgtr \tilde{\mu}_0$.

*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* $n_T = n_I^N < n_I$; *and*
(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 with the preferences of the ISP, so that TMR turns out to be irrelevant: Welfare is the same as under UM since the ISP sets the unregulated market coverage anyway. The regulator makes a difference in case (b) where she imposes neutrality while the ISP would have chosen prioritization under UM. This makes the ISP choose a coverage that is lower than the one that the ISP would have chosen under UM. Rather surprisingly, regulation results in lower coverage and works to the detriment of unserved markets, in the pursuit of providing 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's preferred traffic management practice—prioritization in this case—even though the regulator would have chosen neutrality if she were also in control of coverage. The per-market consumer utility gain that neutrality would bring—which would justify its adoption when there is welfare-maximizing coverage—proves insufficient when there is lower ISP coverage.

### 4.2 Universal Service Obligations (USO)

Under USO, the regulator can choose market coverage $n_U$ 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 $\mu$ and $i = N, P$. 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 prioritization leads to 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 regime 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 that is 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 $n_U = n_*^P < n_*$; and*
(c) *If $\mu > \mu_B$, the ISP chooses P and $n_U = n_*^P = n_*$.*

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

### 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 FB if she imposes both the TMR and the USO regulatory frameworks. We thus assimilate full regulation with the FB benchmark.

### 4.4 Comparisons: Traffic Management and Market Coverage

The comparisons among the three possible regulatory frameworks—FR, TMR, USO—as well as with UM are summarized in Proposition 6 and are illustrated in Fig. 3.

**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 $n_*^N > n_U = n_*^P > n_T = n_I^P \geq n_I^N$, and net neutrality is chosen under FR, while prioritization is chosen under TMR, USO, and UM.*

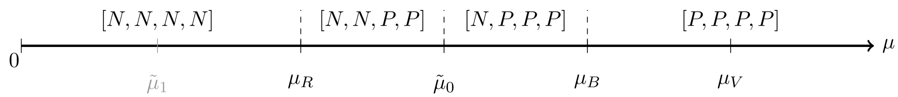


**Fig. 3** Net Neutrality or Prioritization: Outcomes under frameworks [FR,TMR,USO,UM] given bandwidth ($\mu$) availability

(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 all four of the frameworks.*

Figure 3 shows the choice of the traffic management practice *N* or *P* under the three regulatory frameworks FR, TMR, and USO, and under UM, respectively. For example [*P*, *P*, *N*, *N*] means that prioritization is chosen in regulatory frameworks FR and TMR, while net neutrality is chosen under USO and UM.

We see that, whatever is the regulatory framework, net neutrality has a comparative advantage for low bandwidths; and prioritization has the advantage for high bandwidths. Moving from $\mu = 1$ to the right, there exists for each regulatory framework a bandwidth threshold from which prioritization 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.

We see that coverage is always greater when it is chosen by the regulator (FR, USO) than by the ISP (TMR, UM). Note that, from (15) and (16 ), $n_I^i$ and $n_*^i$ are continuous functions of $\mu$. However, as $\mu$ increases, the optimal coverage ($n^U$) exhibits discontinuities. Under FR, when the regulator switches from N to P (for $\mu = \mu^R$), the optimal coverage decreases from $n^U = n_*^N$ to $n^U = n_*^P$. The regulator trades-off a lower coverage with a greater welfare per covered location. As $\mu$ increases further (above $\mu^B$), P is associated with the highest coverage ($n^U = n_*^P > n_*^N$) and the highest benefit per location. Similarly under TMR, the optimal coverage $n^T$ exhibits a discontinuous jump at $\tilde{\mu}_0$ when the regulator switches from N to P. At this cut-off value, P is chosen despite a lower benefit per location because it brings a greater coverage.

### 4.5 The Cost of Incomplete Regulation

From Propositions 3 and 5, it is clear that the FB can be achieved with USO only when there is no disagreement between the regulator and the ISP on the preferred traffic management regime: 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 FB. For the remaining intermediate values of $\mu$, only full regulation can achieve the FB. For this parameter range, we discuss the cost of incomplete regulation[14] and the relative merits of TMR versus USO.

The comparison is done in the following proposition:

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

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 allows the regulator to bring the welfare-maximizing coverage given the $P$ management practice that is chosen by the ISP. However, this management practice is not itself the welfare-maximizing one, so that the result is short of the FB. Similarly, TMR allows the regulator 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.[15]

## 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 assuming that 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. (2002) and Valletti et al. (2002), the question of the choice of the funding mechanism and its effect on the ISP behavior has quickly become a central theme in the literature.[16]

[14] This is an illustration of the classical Tinbergen insight that multiple policy goals cannot generally be achieved with a single instrument.

[15] Note that even in the range $\left(\mu_R, \tilde{\mu}_u\right)$, where welfare is higher under TMR than under USO, USO nevertheless brings 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.

[16] See for instance the contributions on the subject of Chone et al. (2000, 2002) and Gautier and Wauthy (2012).

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 binding. We consider first the case where no compensation mechanism exists. In conformity with the USO literature, we then focus on "self-funded" mechanisms where the ISP losses are "funded through industry subsidies". Note, however, that we analyze the case where funds 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.

**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 the ISP's participation is ensured: whether there exists a range of bandwidth levels for which

$$\Pi^i\left(n^i_*(\mu),\mu\right) = n^i_*(\mu)R^i(\mu) - C\left(n^i_*(\mu),\mu\right) \geq 0, \tag{17}$$

for $i = N, P$. From (1) and (16), this is equivalent to have:

$$R^i(\mu) \geq \frac{1}{2}B^i(\mu). \tag{18}$$

From Lemma 1, we have $\frac{\partial R^i}{\partial \mu} < \frac{\partial B^i}{\partial \mu}$ and $B^i$ tends to infinity when $\mu$ tends to infinity, there exists a maximal bandwidth over which the ISP is not profitable. Since Lemma 1 indicates that $\frac{\partial^2 R^i}{\partial a \partial \mu} < \frac{\partial^2 B^i}{\partial a \partial \mu}$, this maximal bandwidth decreases with $a$. The next proposition uses these facts to define the set $M^i_*$ of bandwidths that satisfy (17).

**Proposition 8** *There exists a $\bar{\mu}^i_*$ such that $\Pi^i\left(n^i_*(\bar{\mu}^i_*),\bar{\mu}^i_*\right) = 0$ and $M^i_* = [1, \bar{\mu}^i_*]$, for $i = N, P$. Moreover $\bar{\mu}^P_* > \bar{\mu}^N_*$.*

Convexity of coverage costs explains the fact that FB coverage proves to be too costly for $\mu > \bar{\mu}^i_*$. Moreover, prioritization is less restrictive on participation than is net neutrality because the ISP can capture a part of the CP's profit. For bandwidth levels that are not in $M^i_*$—for $\mu$ 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 $n^i_\pi$ such that

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

Note that from (1) and (15), this implies that $n^i_\pi = 2n^i_I$. If the participation constraint is binding under both $N$ and $P$, the ISP chooses $N$ under USO if and only if $\mu \leq \mu_R$, while the regulator chooses $N$ 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 affected by the participation constraint, and the results of Proposition 6 still hold.

If $\mu < \bar{\mu}^N_*$, the ISP participation constraint is not binding, and we must return to the relevant case of Proposition 6 with respect to the exact value of $\mu$. If $\mu > \bar{\mu}^P_*$, the ISP participation constraint is always binding and Proposition 6 holds with the modification that $n_U = n^i_\pi$ for all $i$, when regime $i$ applies. Finally, if $\bar{\mu}^N_* < \mu < \bar{\mu}^P_*$, the participation constraint binds under one and only one of the traffic management regime, prioritization as it is less restrictive on participation. Therefore in this case, Proposition 6 is only modified for sub-cases (a) and (b) with $n_U = n^N_\pi$ and remains true.

Note that in this latter case, if traffic management and market coverage are determined independently—possibly by different regulatory authorities, as it is the case in many countries—the lack of coordination between regulators would be detrimental to welfare. Indeed, a regulator that focuses on traffic management would choose N for $\mu \leq \mu^B$ and P otherwise. But in a context where the ISP is budget-constrained, the choice of net neutrality would limit the possibility for the coverage regulator to finance network expansion. In this case, the lack of coordination between the two agencies could lead to a suboptimal choice: a smaller market coverage.

**USO Fund**

Assume now that the regulator is able to establish a USO fund by seizing shares $t_0$ and $t_1$ of the prioritized and non-prioritized CPs' rents, respectively.[17] We let the shares be different for the two CPs as tax avoidance possibilities can in general differ across CPs.[18]

Collecting money from the CPs to finance infrastructure is an argument that is 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 necessarily distort the consumers' demand for content, as does paid prioritization. With a USO fund, the CP could finance infrastructures under NN, and the fund is used to finance infrastructure extension in regions that otherwise would not be covered.

The primary effect of the fund is to enlarge the set $M^i_*$ for which the ISP's is able to supply the first-best coverage. But it can also modify the ISP's behavior. Hereafter, we consider $\mu$ in $M^i_*$ given the existence of the fund and check whether the fund modifies the results under USO and FR.

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

$$\begin{aligned} R^N_t(\mu) &= V^N(\mu) + t_0 a X^N_0(\mu) + t_1 a X^N_1(\mu) > R^N(\mu); \quad \text{and} \\ R^P_t(\mu) &= V^P(\mu) + a\left(1 - t_0\right)\Delta X_0(\mu) + t_0 a X^P_0(\mu) + t_1 a X^P_1(\mu) > R^P(\mu), \end{aligned}$$

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

$$\Delta R_t = \Delta V + a\Delta X_0 + t_1 a \Delta X_1 < \Delta R.$$

[17] An alternative would be to finance the USO fund with money from taxation, as is done in some countries. In this case, the optimal coverage would be reduced as taxation is distortionary.

[18] 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. And large CPs can, for instance, practice tax shifting across countries. Fuchs (2018) gives empirical evidence that some big digital companies intensively employ intangibles that are registered in low tax jurisdictions (as Ireland) and can operate in the market without necessarily being physically present.

Let $n_t^i$ be the profit maximizing coverage under CP taxation. Since, from (15), market coverage is increasing with revenue, we immediately obtain $n_t^i > n_I^i$. Moreover, since $\Delta R_t < \Delta R$, we have that the threshold bandwidth $\mu_t$ that is such that $\Delta R_t = 0$ is greater than $\mu_R$. Prioritization becomes relatively less attractive than does 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 USO 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' rents—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 USO fund does not modify the main results of Sect. 4, except for the fact that neutrality becomes favored by the ISP for a larger bandwidth range.

## 5.2 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 (2016) confirm this conjecture. In this section, we show that these considerations are easily integrated in our model.

Assume that the anonymous CPs that we considered up to now belong to two classes that are different in nature: The CP that is denoted 0 is a large CP (say Google or Facebook); while the CP that is 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 $\Pi_m$ 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 $X = X_0 + X_f$ is the total traffic on the network.

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

$$U\left(X_0, Z_f\left(X_1, ..., X_m\right)\right) = X_0^{\alpha} \cdot Z_f, 0$$

where $Z_f$ is a CES index of the overall fringe consumption that takes into account both sdubstitutability among contents and the number of varieties:[19]

[19] This representation of preferences in a model of monopolistic competition is borrowed from Belleflamme and Peitz (2015, p. 88).

$$Z_f \equiv \left( \sum_{j=1}^{m} X_j^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\beta\sigma}{\sigma-1}} . \tag{19}$$

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 $\beta$ 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: $X_j = \frac{X_f}{m}$. Substituting this value in (19) gives $Z_f = X_f^{\beta} m^{\frac{\beta}{\sigma-1}}$, and the consumer problem can be written as:

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

where $\delta \equiv \frac{\beta}{\sigma-1}$ and where $\omega_0^i$ and $\omega_f^i$ are transmission times of the 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 (5) and (11). Note the case of no fringe diversity that we have considered up to now is a particular case where $m = 1$.

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

$$m^i(n, \mu) = \left( \frac{anX_f^i(\mu)}{F} \right)^{\frac{1}{2}}, \; i = N, P. \tag{20}$$

The fringe aggregate profit is then:

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

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 the demand for fringe content is lower under prioritization than under neutrality, a shift from neutrality to prioritization reduces the 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 reduced demand is less severe as coverage is increased:

**Proposition 9** *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 $V^i = U\big(X_0(\mu), X_f(\mu), m(n, \mu)\big)$ then depends on market coverage:

$$V^N(n,\mu) = n^{-\frac{1}{2}\delta} v(\mu - 1)^{\alpha+\beta+\frac{1}{2}\delta}; \quad \text{and} \tag{21}$$

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

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—is concave in $n$. The functional forms of (21) and (22) with respect to $\mu$ are identical to those of (12) and (13). Results for this version of the model are thus qualitatively similar to those obtained with those in Sects. 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 accompanies greater fringe content demand helps support more variety, which is retroactively valued as such. As both coverage and capacity induce more fringe content consumption, they both contribute to this positive network externality. These effects are, however, less important for priority pricing. As a result, both the ISP and the regulator favor prioritization less often when diversity of content is taken into account.

Note that our results on diversity are in line with those of Reggiani and Valletti (2016) in particular and with the general argument in favor of net neutrality that was first stated by Wu (2003).

### 5.3 Endogenous Bandwidth

Providing enough bandwidth to consumers is an important concern for regulators and there are many projects to increase the minimum bandwidth requirement. In our main analysis, we have considered that the regulator is concerned with the market coverage for a given bandwidth: The objective is to cover as many locations with a given $\mu$. In this section, we discuss the case where the choice of the network capacity is either chosen by the ISP or is regulated.

As we have detailed, an increase in $\mu$ increases consumption ($X$), utility ($V$), revenue ($R$), welfare ($B$), and it twists the preferences towards P. Suppose in this extension that the ISP can invest $F(n)$ to increase bandwidth from $\mu$ to $\mu'$.

The incentives of the regulator and the ISP to increase the bandwidth under $i = N, P$ can be measured respectively by $n\left(B^i(\mu') - B^i(\mu)\right)$ and $n\left(R^i(\mu') - R^i(\mu)\right)$. From Lemma 1, we know that the former is greater than the latter; this implies that the preferences of the ISP and the regulator may not be aligned. If, under $i = N, P$, $n\left(B^i(\mu') - B^i(\mu)\right) \geq F(n) \geq n\left(R^i(\mu') - R^i(\mu)\right)$, then the ISP will restrict the bandwidth compared to the welfare maximizing level. Hence, there is room for regulatory intervention, and bandwidth regulation could be part of the USO.

The capacity improvement may change the preferred traffic management regime: Under TMR and FR, if $\mu' > \mu^B \geq \mu$, the traffic regulation changes to P. This also occurs under USO and UM if $\mu' > \mu^R \geq \mu$. Finally notice that the capacity extension will call for a higher coverage as $n_I^i(\mu)$ and $n_*^i(\mu)$ are both increasing in $\mu$. In that sense, bandwidth and coverage are complements.[20]

## 6 Conclusion

In order to assess what is the most efficient regulation to reach the common goal of a universal broadband coverage in a context where moving from net neutrality to prioritization is on the agenda, we have considered three regulatory frameworks: traffic management regulation (TMR); universal service obligations (USOs); and full regulation (FR).

We have assessed that, 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 unregulated 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 is less inclined to move from net neutrality to prioritization than is the ISP: Prioritization can be forbidden even though the ISP would adopt it. For relatively low or relatively high bandwidth levels, imposing TMR is redundant, since the interests of the ISP and the regulator converge for net neutrality in the event of narrow bandwidth, and for prioritization in the event of large bandwidth.

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

So, in a range of intermediate values of bandwidth, FR is required to attain maximum welfare because USO regulation 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 of bandwidth, while the reverse is true for higher values of bandwidth. 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 USOs, and conversely for higher bandwidth levels. But globally, universal service obligations appear to be a stronger regulatory instrument than the imposition of traffic management regulation since USOs reach first-best for

[20] Notice that if the ISP does not find it profitable to invest in bandwidth expansion, imposing such an obligation reduces the ISP's revenue, and this may therefore limits its possibility to finance network extension. In the case of a budget-constrained ISP, imposing bandwidth obligations conflicts with coverage obligations, and the regulator must choose between greater coverage and lower bandwidth: covering more locations with a relatively lower $\mu$, or fewer locations with a higher $\mu$.

low and high bandwidth levels, while TMR never reaches maximal welfare, and it is only for a subset of intermediate values where USOs fail to reach first-best that TMR is welfare-superior. It is important to note, however, that if USOs lead to the first-best for high bandwidth, it is because our definition of USOs allows the ISP to choose the traffic regime. In practice, in most countries, 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 the internet. In the 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 (2003). Net neutrality contributed to the universal adoption of the internet and to the diversity of content that it delivered, with an important increase of content that required expanded broadband in the early 2010's.

But, in a certain sense, net neutrality became a victim of its own success as it fed debates to its economic efficiency when concerns over congestion and misallocation of traffic due to equal treatment of contents of different time sensitivities arose (Peitz and Schuett, 2016). In terms of regulation, such debates lead to a state of flux, as exemplified by the different rulemakings of the FCC from 2015. If we consider that bandwidth requirements are currently at an intermediate stage of growth, our model is in line with both choices of maintaining or repealing NN, but it suggests that there will be a tendency, in the long-term, to repeal net neutrality.

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 market coverage adds a trade-off between the extensive margin and the intensive margin. This puts the debate on NN into a better perspective as a change in the extensive margin has more effect on the network benefits than does a change in 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 that have been studied in either one or both literature strands. The most important limitation of our model is the assumption 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., 2002)), for instance) and the analysis, in the NN literature, of CPs that are able to invest in their own infrastructure so as to improve their quality of service ((Choi et al., 2018), for instance).

**Table 2** Nomenclature of variables

| Index | Description |
|---|---|
| $a$ | Expected benefit per unit of traffic |
| $\alpha$ | Preference parameter for content 0 |
| $\beta$ | Preference parameter for content 1 |
| $B^i$ | Per location social benefit function in regime $i$ |
| $C$ | Cost function |
| $c$ | Scale parameter for the cost function |
| $\delta$ | Composite preference parameter (only in 5.2) |
| $\Delta H$ | Variation $H^P - H^N$ for any function $H$ |
| $F$ | Fringe fixed entry cost |
| $i$ | Traffic management practice index, N or P |
| $j$ | Decision maker index, I for ISP or * for Regulator |
| $\lambda$ | Frequency per second of data transmission |
| $m$ | Number of small CPs in the fringe (only in 5.2) |
| $n$ | Location (any) |
| $n_j^i$ | Optimal coverage in regime $i$ and for decision maker $j$ |
| $N$ | Index for Net-Neutrality |
| $p$ | Users' connection fixed charge |
| $P$ | Index for Prioritization |
| $\Pi^i$ | Total ISP profit function in regime $i$ |
| $R^i$ | Per location ISP profit function in regime $i$ |
| $\rho$ | Prioritization traffic parameter |
| $t_h$ | Seizing shares of content $h = 0, 1$ (only in 5.1) |
| $V^i$ | Per location indirect utility function in regime $i$ |
| $U$ | Utility function for any consumer per location |
| $W^i$ | Total welfare function in regime $i$ |
| $X_h$ | Total traffic per location of CP $h = 0, 1$ |
| $X$ | Total traffic per location of all CPs |
| $Z_f$ | CES index of the overall fringe consumption (only in 5.2) |
| $\bar{\omega}$ | Average waiting time |
| $\omega_h$ | Waiting time for CP $h = 0, 1$ |

## Appendix

**Nomenclature of variables** We provide a nomenclature of variables that are used in our model. Depending on the regimes that are studied all of them can be stared, tilded or hatted and adorned with a subscript or superscript letter/number (see Table 2).

***Proof of Lemma 1***

(a) Directly for Net Neutrality, $B^N(\mu) - R^N(\mu) = aX^N(\mu) > 0$. For Prioritization, $B^P(\mu) - R^P(\mu) = a\big(X_1^P(\mu) + X_0^N(\mu)\big) > 0$.

(b) Let us define $v^i(\mu) := \frac{\partial V^i(\mu)}{\partial \mu}$, $r^i(\mu) := \frac{\partial R^i(\mu)}{\partial \mu}$ and $b^i(\mu) := \frac{\partial B^i(\mu)}{\partial \mu}$. With Net Neutrality we have

$$v^N(\mu) = (\alpha+\beta)v(\mu-1)^{\alpha+\beta-1} = \frac{\alpha+\beta}{\mu-1}V^N(\mu) > 0,$$

so $v^N(\mu)$ decreases with $\mu$. Moreover $r^N(\mu) = v^N(\mu)$ and $b^N(\mu) = v^N(\mu) + a$, which decreases with $\mu$, and $b^N(\mu) > r^N(\mu)$ for all $\mu \geq 1$. With Prioritization,

$$v^P(\mu) = \frac{H(\mu)}{\mu(\mu-1)((1-\rho)\mu+\rho)}V^P(\mu) > 0,$$

where $H(\mu) = (2\alpha+\beta)(1-\rho)\mu^2 - \alpha(1-2\rho)\mu + \beta\rho > 0$, as $H(1) = \alpha+\beta > 0$; $H'(1) = 2(1-\rho)\beta + (3-2\rho)\alpha > 0$ and $H''(\mu) = (2\alpha+\beta)(1-\rho) > 0$. Consequently

$$r^P(\mu) = v^P(\mu) + a\frac{\partial \Delta X_0(\mu)}{\partial \mu} \text{ with } \frac{\partial \Delta X_0(\mu)}{\partial \mu} = \frac{2\alpha}{(\alpha+\beta)}(\mu-1)(1-\rho) > 0;$$

and

$$b^P(\mu) = v^P(\mu) + a\frac{\partial X^P(\mu)}{\partial \mu}$$

$$\text{with } \frac{\partial X^P(\mu)}{\partial \mu} = \frac{\rho\beta}{\mu^2(\alpha+\beta)} + \frac{\alpha+\beta(1-\rho)}{\alpha+\beta} + \frac{\partial \Delta X_0(\mu)}{\partial \mu} > \frac{\partial \Delta X_0(\mu)}{\partial \mu}.$$

Indeed $b^P(\mu) > r^P(\mu)$ for all $\mu \geq 1$, as

$$b^P(\mu) - r^P(\mu) = a\left(\frac{\rho\beta}{\mu^2(\alpha+\beta)} + \frac{\alpha+\beta(1-\rho)}{\alpha+\beta}\right) > 0.$$

(c) We have

$$\frac{\partial^2 R^N(\mu)}{\partial a \partial \mu} = 0 < \frac{\partial^2 B^N(\mu)}{\partial a \partial \mu} = 1,$$

and

$$0 < \frac{\partial^2 R^P(\mu)}{\partial a \partial \mu} = \frac{\partial \Delta X_0(\mu)}{\partial \mu} < \frac{\partial^2 B^N(\mu)}{\partial a \partial \mu} = \frac{\partial X^P(\mu)}{\partial \mu}.$$

***Proof of Lemma 2***

(a) From (5) and (11), we form

$\Delta X \equiv X^P(\mu) - X^N(\mu) = \Delta X \equiv X^P(\mu) - X^N(\mu),$

which writes

$$\Delta X = (\alpha(1-\rho)\mu - \beta\rho)\frac{(\mu-1)^2}{(\alpha+\beta)\mu} \lesseqqgtr 0 \text{ as } \mu \lesseqqgtr \mu_X \equiv \frac{\beta}{\alpha}\frac{\rho}{1-\rho}, \tag{23}$$

with $\mu_X > \frac{\rho}{1-\rho}$ as $\beta > \alpha$. Note that $\mu_X > 1$ iff $\frac{\beta}{\alpha+\beta} > \rho$. So if $\rho \geq \frac{\beta}{\alpha+\beta}$ then $\Delta X \geq 0$ for all $\mu > 1$.

(b) From (12) and (13), we form

$$\frac{V^P}{V^N} = \frac{((1-\rho)\mu+\rho)^{\alpha+\beta}}{\mu^\beta},$$

so that

$$V^P \lesseqqgtr V^N \text{ as } G(\mu) \equiv \frac{\ln((1-\rho)\mu+\rho)}{\ln\mu - \ln((1-\rho)\mu+\rho)} \lesseqqgtr \frac{\beta}{\alpha}. \tag{24}$$

Indeed, $G(\mu)$ is well defined as for all $\mu > 1$ and $\rho \in ]0,1[$ we have $\mu > (1-\rho)\mu + \rho > 1$. Note that one can also write $G(\mu) = (g(\mu)-1)^{-1}$, where $g(\mu) = \frac{\ln\mu}{\ln((1-\rho)\mu+\rho)} > 1$ is decreasing and convex with respect to $\mu > 1$. As $G'(\mu) = -G^2(\mu)g'(\mu) > 0$ then $G(\mu)$ is a strictly increasing and concave function of $\mu$ such that $\lim_{\mu\to 1} G(\mu) = \frac{1-\rho}{\rho}$ and $\lim_{\mu\to\infty} G(\mu) = \infty$. So, if $\frac{\beta}{\alpha+\beta} > \rho$, there exists $\mu_V : G(\mu_V) = \frac{\beta}{\alpha} \Leftrightarrow \frac{\rho}{1-\rho}G(\mu_V) = \mu_X > 1$. Moreover as $\lim_{\mu\to 1} G'(\mu) = -\left(\frac{1-\rho}{\rho}\right)^2\left(-\frac{1}{2}\frac{\rho}{1-\rho}\right) < \frac{1-\rho}{\rho}$ and by concavity of $G(\mu)$, we have

$$1 \leq \frac{\rho}{1-\rho}G(\mu) \leq \mu \text{ for } \mu \geq 1. \tag{25}$$

This implies that $\mu_V > \frac{\rho}{1-\rho}G(\mu_V) = \mu_X$.

(c) From (23) we see straightforwardly that $\frac{\partial\mu_X}{\partial\rho} > 0$. Differentiating $G(\mu_V) = \frac{\beta}{\alpha}$ with respect to $\rho$ yields $G'(\mu_V)\frac{\partial\mu_V}{\partial\rho} = -\frac{\partial G(\mu)}{\partial\rho}$. As $\frac{\partial G(\mu)}{\partial\rho} = \frac{(\mu-1)\ln\mu}{((1-\rho)\mu+\rho)(\ln((1-\rho)\mu+\rho)-\ln\mu)^2} < 0$; this implies $\frac{\partial\mu_V}{\partial\rho} > 0$.

***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 $\Delta X_1(\mu) < 0$. When $\alpha < \beta$, 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, $\forall\mu$, there exists a $\mu_R < \mu_V$ such that $\Delta R(\mu) = \Delta V(\mu) + a\Delta X_0(\mu) \lesseqqgtr 0$ as $\mu \lesseqqgtr \mu_R$. As a result $\Delta R(\mu)$ is locally increasing around $\mu = \mu_R : \frac{\partial\Delta R(\mu_R)}{\partial\mu} > 0$, then $\frac{\partial\mu_R}{\partial\rho} = -\frac{\partial\Delta R(\mu)}{\partial\rho}\Big/\frac{\partial\Delta R(\mu)}{\partial\mu} > 0$ since $\frac{\partial\Delta R(\mu)}{\partial\rho} = \frac{\partial\Delta V(\mu)}{\partial\rho} + a\frac{\partial\Delta X_0(\mu)}{\partial\rho} < 0$ because

$$\frac{\partial \Delta X_0(\mu)}{\partial \rho} = -\frac{\alpha}{\alpha+\beta}(\mu-1)^2 < 0; \quad \text{and}$$
$$\frac{\partial \Delta V(\mu)}{\partial \rho} = -(\alpha+\beta)(\mu-1)((1-\rho)\mu+\rho)^{\alpha+\beta-1}V^N(\mu) < 0.$$

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 $\Delta B(\mu) < 0 = \Delta R(\mu)$ at $\mu = \mu_R$, there exists a $\mu_B > \mu_R$ such that $\Delta B(\mu) \lesseqqgtr 0$ 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 $\mu_B < \mu_V$. With the use of same arguments as above, $\frac{\partial \mu_B}{\partial \rho} = -\frac{\partial \Delta B(\mu)}{\partial \rho} Big/ \frac{\partial \Delta B(\mu)}{\partial \mu} > 0$ as $\frac{\partial \Delta X(\mu)}{\partial \rho} = -(\alpha\mu+\beta) < 0$. □

***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 $\Pi^N$ 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 $\mu$, by definitions (15) and (16) 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 $\mu$, coverage solutions are given by $N(x,\mu) = \frac{x}{c\mu}$, which is an increasing function in $x$. So we have:

$$\begin{aligned} &W^P(n_I^P(\mu), \mu) - W^N(n_I^N(\mu), \mu) \\ &\quad = N(R^P(\mu), \mu)B^P(\mu) \\ &\qquad - C(n_I^P(\mu), \mu) - N(R^N(\mu), \mu)B^N(\mu) + C(n_I^N(\mu), \mu). \end{aligned}$$

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

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

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:

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

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 $W^P(n_I^N(\tilde{\mu}_0), \tilde{\mu}_0) = W^N(n_I^P(\tilde{\mu}_0), \tilde{\mu}_0)$. 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

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

This proves that $W^P(n_I^P(\mu), \mu) - W^N(n_I^N(\mu), \mu) < 0$ 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

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

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

***Proof of Proposition 5*** At the second stage, the ISP chooses the regime independently of coverage, so that the regime is $N$ if $\mu \le \mu_R$ and $P$ if $\mu > \mu_R$. If $\mu \le \mu_R$ or $\mu > \mu_B$, the regime chosen by the ISP is also the regime that is 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

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

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

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

so that $N$ is chosen under FR.

(b) if $\tilde{\mu}_1 \le \mu < \mu_R$ then $B^P(\mu) \ge R^N(\mu)$ so that $n_*^P \ge n_I^N$. 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 $\mu \in \left(\tilde{\mu}_0, \mu_B\right)$, 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 $\mu = \tilde{\mu}_0 : W^P\left(n_*^P\left(\tilde{\mu}_0\right), \tilde{\mu}_0\right) > W^P\left(n_I^P\left(\tilde{\mu}_0\right), \tilde{\mu}_0\right) = W^N\left(n_I^P\left(\tilde{\mu}_0\right), \tilde{\mu}_0\right) > W^N\left(n_I^N\left(\tilde{\mu}_0\right), \tilde{\mu}_0\right)$ Now we have

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

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

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

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 $W^P\left(n_*^P\left(\tilde{\mu}_u\right), \tilde{\mu}_u\right) = W^N\left(n_I^N\left(\tilde{\mu}_u\right), \tilde{\mu}_u\right)$ and $W^P\left(n_*^P(\mu), \mu\right) \lesseqgtr W^N\left(n_I^N(\mu), \mu\right)$ as $\mu \lesseqgtr \tilde{\mu}_u$. □

***Proof of Proposition 8*** *First*, let us give a characterization of $M_*^i$. Inserting first-best coverages $n_*^i(\mu) = \frac{B^i(\mu)}{c\mu}$ into the ISP participation constraint (17) implies that $\Pi^i\left(n_*^i(\mu), \mu\right) = R^i(\mu) - \frac{1}{2}B^i(\mu) \geq 0$. This can also be written as

$\frac{1}{2}\left(V^i(\mu) - aZ^i(\mu)\right) \geq 0$, where $Z^N(\mu) = X^N(\mu)$ and $Z^P(\mu) = X^P(\mu) - 2X_0^N(\mu) = \frac{1}{2}(\mu - 1)\frac{\alpha\mu^2+(\beta-3\alpha)\mu+\beta}{\mu(\alpha+\beta)} > 0$.[21] Here $Z^N(\mu)$ is increasing and linear, and $Z^P(\mu)$ is increasing and concave for all $\mu > 1$. Then for $a = 0$, we have $\frac{1}{2}V^i(\mu) - aZ^i(\mu) > 0$ and

$$\lim_{a\to+\infty} \left(V^i(\mu) - aZ^i(\mu)\right) = -Z^i(\mu) \lim_{a\to+\infty} (a) = -\infty.$$

So there exists a unique $a_\pi^i : \frac{1}{2}\left(V^i(\mu) - aZ^i(\mu)\right) = 0$. More precisely

$$a_\pi^N = \frac{V^N(\mu)}{X^N(\mu)} = \nu(\mu - 1)^{\alpha+\beta-1} > 0; \quad \text{and}$$
$$a_\pi^P = \frac{V^P(\mu)}{Z^P(\mu)} = \nu(\alpha + \beta)\left(\frac{1}{2}\right)^{\alpha+\beta-1} \frac{(\mu - 1)^{\alpha+\beta-1}(\mu + 1)^{\alpha+\beta}\mu^{-(1+\beta)}}{\alpha\mu^2 + (\beta - 3\alpha)\mu + \beta} > 0.$$

When $\alpha + \beta < 1$, these price thresholds are decreasing functions of $\mu$ with $\lim_{\mu\to+\infty} a_\pi^i = 0$, so that for all $a$, there exists a unique $\bar{\mu}_*^i : a_\pi^i = a$. Then $\bar{\mu}_*^i$ becomes the inverse function of $a_\pi^i$, defined from $\mathbb{R}_+$ to $[1, +\infty)$ and it is decreasing in $a$. As a result the ISP participation constraint (17) is satisfied for $\mu \leq \bar{\mu}_*^i$ and $M_*^i = [1, \bar{\mu}_*^i]$. Note that for a given $\mu$

$$a_\pi^P < a_\pi^N,$$

as for all $\mu > 1$ we have

$$(\alpha + \beta)\left(\frac{1}{2}\right)^{\alpha+\beta-1} \frac{(\mu + 1)^{\alpha+\beta}\mu^{-(1+\beta)}}{\alpha\mu^2 + (\beta - 3\alpha)\mu + \beta} < \frac{1}{\mu^2} < 1.$$

This implies that $\bar{\mu}_*^N$ and $\bar{\mu}_*^P$ cannot intersect for any $a$.

*Second*, define just for the purpose $\bar{\mu}_*^i$ as functions of $a$, and let us define the following ad-price threshold

$$a_B^P : \bar{\mu}_*^P(a_B^P) = \mu_B.$$

It exists and is unique as $\bar{\mu}_*^P(a)$ is a decreasing one-to-one mapping from $\mathbb{R}_+$ to $[1, +\infty)$, but $\mu_B$ is from $\mathbb{R}_+$ to $[\mu_X, \mu_V]$. Then by definition, we have, when $a = a_B^P : n_*^P\left(\bar{\mu}_*^P(a_B^P)\right) = n_*^N\left(\mu_B\right) = n_*^P\left(\mu_B\right)$ and $\Pi^N\left(n_*^N\left(\mu_B\right)\right), \mu_B\right) < \Pi^P\left(n_*^N\left(\mu_B\right)\right), \mu_B\right) = \Pi^P\left(n_*^P\left(\mu_B\right)\right), \mu_B\right) = 0$ as, using Proposition 1, we have that for all $(n, \mu) : \Pi^P(n, \mu) > \Pi^N(n, \mu)$ if $\mu > \mu_R$, as $\Delta\Pi = n\Delta R$. Here, $\mu = \mu_B > \mu_R$, so the ISP participation constraint for $i = N$ is violated: this implies that $\bar{\mu}_*^P(a_B^P) > \bar{\mu}_*^N(a_B^P)$. As $\bar{\mu}_*^N(a)$ and $\bar{\mu}_*^P(a)$ cannot intersect, this proves that $\bar{\mu}_*^N < \bar{\mu}_*^P$. □

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

***Proof of Proposition 9*** From (5), (11) and ( 20), we obtain:

$$m'_n = \frac{1}{2}\left(\frac{aX_f^i}{F}\right)^{\frac{1}{2}} n^{-\frac{1}{2}} > 0;$$
$$m'_\mu = \frac{1}{2}\left(\frac{an}{F}\right)^{\frac{1}{2}} \left(X_f^i\right)^{-\frac{1}{2}} \left(X_f^i\right)' > 0; \quad \text{and}$$
$$m''_{\mu n} = \frac{1}{4}\left(\frac{an}{F}\right)^{-\frac{1}{2}} \left(X_f^i\right)^{-\frac{1}{2}} \left(X_f^i\right)' > 0.$$

Moreover, since

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

we obtain:

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

□

**Acknowledgements** We thank all participants at the MRE seminar at University of Montpellier, at the 61th Congress of the *Société canadienne de science économique* (Montréal, Québec), at the 49th Annual Conference of the *European Association for Research in Industrial Economics* (Vienna, Austria), and at the 9th *Summer School of Association Francophone de Recherche en Economie Numérique* (Avignon, France) and the editor L. White and the referees for their useful comments and suggestions. All remaining errors are our own.

## References

Anton, J., Vander Weide, J. H., & Vettas, N. (2002). Entry auctions and strategic behavior under cross-market price constraints. *International Journal of Industrial Organization, 20*, 611–629.

Belleflamme, P., & Peitz, M. (2015). *Industrial Organization*. Cambridge: Cambridge University Press.

Bennet, M. F., King, A. S. Jr., Machin III, J. & Porter, R. (2021) Open Letter, available through https://www.bennet.senate.gov/public/index.cfm/2021/3/bennet-king-portman-manchin-urge-biden-administration-to-create-modern-unified-federal-broadband-standard. Accessed 2 Jan 2024.

Bourreau, M., Feasy, R. & S. Hoernig, (2017). Demand-Side Policies to Accelerate the Transition to Ultrafast Broadband. Centre on Regulation in Europe (CERRE).

Bourreau, M., Kourandi, F., & Valletti, T. (2015). Net neutrality with competing internet platforms. *The Journal of Industrial Economics, LXIII*, 30–73.

BroadbandNow. https://broadbandnow.com/report/fcc-broadband-definition/. Accessed 2nd Jan 2024.
Canadian Radio-television and Telecommunications Commissions. (2021). Broadband Fund: Closing the digital divide in Canada, available through https://crtc.gc.ca/eng/internet/internet.htm. Accessed 2nd Jan 2024.
Choi, J. P., Jeon, D.-S., & Kim, B.-C. (2015). Net Neutrality. *Business Models, and Internet Interconnection, American Economic Journal: Microeconomics, 7*, 104–141.
Choi, J. P., Jeon, D.-S., & Kim, B.-C. (2018). Net Neutrality, Network Capacity, and Innovation at the Edges. *The Journal of Industrial Economics, LXVI*, 172–204.
Choi, J. P., & Kim, B.-C. (2010). Net Neutrality and Investment Incentives. *Rand Journal of Economics, 41*, 446–471.
Chone, P., Flochel, L., & Perrot, A. (2000). Universal Service Obligations and Competition. *Information Economics and Policy, 12*, 249–259.
Chone, P., Flochel, L., & Perrot, A. (2002). Allocating and Funding Universal Service Obligations in a Competitive Market. *International Journal of Industrial Organization, 20*, 1247–1276.
Congressional Research Service. (2023). FCC Adopts Proposed Net Neutrality Rule, Available through https://crsreports.congress.gov/product/pdf/IF/IF12513.
Economides, N., & Hermalin, B. E. (2012). The economics of network neutrality. *Rand Journal of Economics, 43*, 602–629.
FCC, 2023. FCC Launches Inquiry to Increase Minimum Broadband Speed Benchmark, Set Gigabit Future Goal, News Release available through https://www.fcc.gov/document/fcc-launches-inquiry-increase-minimum-broadband-speed-benchmark-0. Accessed 2 Jan. 2024.
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.
Garcia-Calvo, A. (2012). Universal Service Policies in the Context of National Broadband Plans, OECD Digital Economy Papers. No., *203*, 2024. https://doi.org/10.1787/5k94gz19flq4-en. Accessed 2 Jan
Gautier, A., & Wauthy, X. (2010). Price Competition under Universal Service Obligations. *International Journal of Economic Theory, 6*, 311–326.
Gautier, A., & Wauthy, X. (2012). Competitively neutral universal service obligations. *Information Economics and Policy, 24*, 254–261.
Katz, M. L. (2017). Whither U.S. net neutrality regulation? *Review of Industrial Organization., 50*, 441–468.
Krämer, J., & Wiewiorra, L. (2012). Network neutrality and congestion sensitive content providers. *Information Systems Research, 23*, 1303–1321.
Krämer, J., Wiewiorra, L., & Weinhardt, C. (2013). Net neutrality: A progress report. *Telecommunications Policy, 37*, 794–813.
McMenemy, D. (2022). Internet access and bridging the digital divide: The crucial role of universal service obligations in telecom policy. In M. Smiths (Ed.), *Information for a Better World: Shaping the Global Future. Series: Lecture Notes in Computer Science, 13192* (pp. 122–134). Cham: Springer.
Peitz, M., & Schuett, F. (2016). Net neutrality and inflation of traffic. *International Journal of Industrial Organization, 46*, 16–62.
Poudou, J.-C., & Roland, M. (2014). Efficiency of uniform pricing in universal service obligations. *International Journal of Industrial Organization, 37*, 141–152.
Poudou, J.-C., & Roland, M. (2017). Equity justifications for universal service obligations. *International Journal of Industrial Organization, 52*, 63–95.
Reggiani, C., & Valletti, T. (2016). Net neutrality and innovation at the core and at the edge. *International Journal of Industrial Organization, 45*, 16–27.
Valletti, T., Hoernig, S., & Barros, P. (2002). Universal service and entry: The role of uniform pricing and coverage constraints. *Journal of Regulatory Economics, 21*, 169–190.
Wu, T. (2003). Network neutrality, broadband discrimination. *Journal of Telecommunications and High Technology, 2*, 141–179.