Journal of Economic Behavior and Organization 191 (2021) 620–638

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

An economic model of the last-mile internet[☆]

Rakesh Chaturvedi[a,*], Souvik Dutta[b], Kiriti Kanjilal[b]

[a] *B-205, R & D Block, Indraprastha Institute of Information Technology Delhi, Okhla Industrial Estate, Phase 3, New Delhi 110020, India*
[b] *Indraprastha Institute of Information Technology Delhi, India*

ARTICLE INFO

Article history:
Received 11 September 2020
Revised 21 July 2021
Accepted 18 September 2021

JEL classification:
D42
D43
D46
L11
L12
L13
L86
L96

Keywords:
Internet
Network neutrality
Congestion pricing
Two-part tariffs
Nonlinear pricing

ABSTRACT

Pricing decisions of an internet service provider (ISP) are studied in a model based on complementarity between broadband connection and content, congestion externalities on consumer side and oligopolistic externalities on content provider side. When consumers face two-part tariffs from the ISP, the equilibrium is sensitive to usage price level but is invariant to its structure on two sides. With nonlinear pricing, the markup of content providers affects consumer prices while congestion externalities and elasticity of content demand shape the price for providers. For the zero-price rule, a neutrality-of-policy result holds with two-part tariffs but not with nonlinear pricing.

1. Introduction

The commercial internet supports a wide range of activities in domains like business, media, entertainment and education. Structurally, the global internet is a network of networks. The supporting physical infrastructure is described in detail in the recent surveys of Greenstein (2020) and Economides (2007). Questions about pricing and investment incentives to improve capacity have been of central interest since the transition of the internet from a government sponsored project to a private enterprise. This is because streaming and gaming applications, comprising the majority of internet traffic,[1] are data-intensive; moreover, they lose their value if users experience delay in these services. So efficient congestion management is an important issue on the internet. A public policy issue related to pricing and investment that has received the

☆ This research did not receive any specific grant from funding agencies from the public, commercial, or not-for-profit sectors.

* Corresponding author.
E-mail addresses: rakesh@iiitd.ac.in (R. Chaturvedi), souvik@iiitd.ac.in (S. Dutta), kanjilal@iiitd.ac.in (K. Kanjilal).

[1] Globally, the largest share of internet usage is for video streaming. However, file sharing, cloud services, gaming etc have also gained substantial marketshare recently. Cloud services are not only important for storage, but power other services like video gaming, certain software etc. (Source: Sandvine Global Internet Phenomena Report 2020). Specifically, the public cloud market revenue worldwide is set to increase from 26.4 Billion USD in 2012 to a projected 552 billion USD in 2027 (Source: https://www.statista.com/statistics/477702/public-cloud-vendor-revenue-forecast/)

https://doi.org/10.1016/j.jebo.2021.09.025
0167-2681/

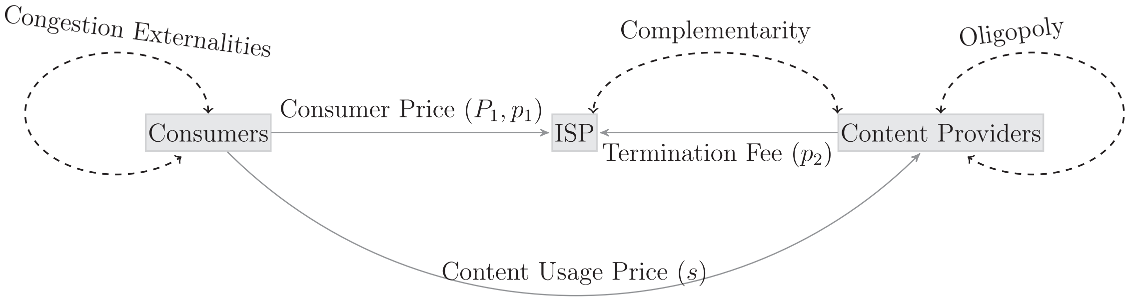

Fig. 1. Model Structure. The acronym ISP stands for internet service provider.The entities in the model are in the boxes. The solid connections indicate the flow of money. The dashed connections point out the salient economic features.

most attention in regulatory as well as in the recent academic literature is that of network neutrality. A term coined by Wu (2003), network neutrality is broadly understood to mean non-discriminatory treatment of all data packets flowing on service provider's communication channels. This means that network management practices like introducing priority lanes[2] for certain data packets and differential pricing treatment of certain data packets are prohibited.

While there are multiple facets of the network neutrality debate, a crucial one is what (Hemphill, 2008) refers to as the zero-price rule that prohibits last mile internet service providers from charging any *termination*[3] fees to non-originating content providers[4] for access to their customer base. A regulatory move that does away with zero-price rule would introduce the prospect of two-sided pricing for the internet service providers wherein they would be able to price both the consumer side and the provider side. Lee and Wu (2009) discuss this aspect of network neutrality at length and the models in this paper help evaluate the desirability of a zero-price rule.

In this paper we present a formal evaluation of pricing and investment policies in a simple model of the internet ecosystem. Our model has a continuum of consumers who consume content, are heterogenous in taste, and suffer congestion externalities; an oligopolistic market populated by content providers who supply content in a Cournot supply subgame; and a monopoly service provider which supplies broadband capacity and sets prices for consumers and for content providers. The primary economic aspect of the internet ecosystem that we take as fundamental to modeling is the complementarity of broadband connection and content for the consumers.[5] The strategic interaction between the broadband service provider and the content providers is modeled as a variant of a Cournot game of pricing complementary goods. The primary contribution from a modeling standpoint is casting the internet in a demand-supply framework that is achieved by a coherent integration of Cournot's two models - competitive supply model of substitute goods and competitive pricing model of complementary goods. Our model, compactly depicted in Fig. 1, presents a simple framework to study pricing of data flows on the internet and investment incentives for the service provider.

We first study a baseline model in which consumers who are heterogenous in tastes face two-part tariffs and content providers face linear termination charges. Our first result gives a striking conclusion that the service provider's equilibrium profit depends only on the usage price level on the network; for any given usage price level, it is invariant to the usage price structure, that is, to the way it is split between the consumer side and the content provider side. The usage price level is given by a Lerner-style formula in which the markup is over the *Pigouvian* tax that a social planner would set. This result leads to the conclusion that the network neutrality regulation of zero-price rule has no impact on the equilibrium other than forcing a change in usage price for consumers that keeps the price level unchanged. This is a neutrality-of-policy result that generalizes a similar result obtained in Greenstein et al. (2016) in a very simple[6] expository model. The neutrality result prompts us to extend our baseline model to one in which consumers face nonlinear prices from the service provider. We derive equilibrium prices on both sides of the market. The equilibrium termination fee is given by a Lerner-style formula that takes into account its impact on network congestion and the content market. The equilibrium nonlinear price schedule on the consumer side is completely pinned down by the usage allocation rule. This pegging of pricing rule onto allocation rule is standard fare in classic screening models. What is interesting is that under some conditions on distributions of tastes, the optimal allocation for a consumer of any type is found by equating the virtual marginal benefit net of content providers' margin to the service provider's marginal cost. In other words, the complementarity relationship forces the monopolist to

[2] Lane prioritization issues are not the focus of our work here. Pil Choi and Kim (2010) and Economides and Hermalin (2012) are among papers that deal with these issues.

[3] This terminology is inherited from its usage in the telephone networks wherein network A would charge network B for calls that originate from network B and terminate in network A.

[4] Content providers do pay to a service provider for their own internet connection. By non-originating content providers, we mean those who do not have a direct contract for internet connection with the service provider in question.

[5] The emphasis on complementarity is also shared by Greenstein et al. (2016) in their survey who use Cournot (1838) model of pricing complementary goods as their workhorse model to exposit many insights.

[6] The neutrality result in Greenstein et al. (2016) is obtained in a model with a single service provider, single content provider, only usage prices on both sides, no taste heterogeneity among consumers and a demand function that is assumed without modeling consumer preferences.

cede some economic pie (the size of which is given by the margin) to the content providers in addition to ceding some of it to the consumers on account of asymmetric information.

We also analyze welfare properties of equilibrium in both versions of the model. In the baseline model with two-part tariffs, the equilibrium usage price level is too low relative to the socially optimal level if and only if the marginal consumer demand is bigger then the average consumer demand. This result is reminiscent of classic welfare analysis of two-part tariffs set by a monopolist. In the extended model with nonlinear pricing, equilibrium price on the consumer side is always too high relative to the social optimum. On the provider side, the equilibrium termination fee is socially optimal. Our last result relates the welfare effects of a zero-price rule to characteristics of consumer demand for content. In particular, a positive termination fee is socially suboptimal if the consumer demand for content is sufficiently inelastic. However, when the demand is elastic, then a zero termination fee is socially suboptimal. In both the baseline and the extended model, absent any regulatory intervention in prices, the equilibrium investment decision is welfare optimal while the consumer market size is too small relative to the social optimum.

There is a large literature on many aspects of network neutrality and we make no attempt to summarize it here. Readers are directed to a recent survey by Greenstein et al. (2016) and references therein. Here, we clarify the relationship of our model to the literature on two-sided markets spawned by Rochet and Tirole (2003, 2006) among others. Although we use some of the same language (there is a consumer side and a provider side), there are important differences. First, the network effects (Farrell and Saloner, 1985; Farrell and Saloner, 1986 and Katz and Shapiro, 1985; Katz and Shapiro, 1986) or cross group positive externalities that are characteristic of these markets may or may not obtain. In our model, a consumer wants more providers if competition can drive down the price of content; otherwise not. Similarly, a provider cares about more consumers if he can look forward to more sales; otherwise not. In other words, the nature of cross group externalities is endogenous. Second, there are negative usage externalities within each side. Consumers exert congestion externalities among themselves while providers exert externalities that are intrinsic to oligopoly. Third, one of the sides here comprises of profit-maximizing content provider firms. Consequently, competitive forces rather than service provider's tariffs drive the equilibrium number of providers in the market. Fourth, the added layer is that the two sides cannot transact without the platform; moreover, one of the sides (provider side) can price the usage decisions of the other (consumer side). A fifth salient difference is the accounting of volume of interactions given memberships. Rochet and Tirole (2006) primarily use the multiplicative volume assumption which means that the number of interactions on the platform is simply the product of the number of entities on both sides. In our model, the interactions of interest are data flows. So the volume of interactions on the platform will simply be the total content flow through the service provider's communication channels which is endogenously determined. Similar to our work, Economides and Tåg (2012) also focus on the welfare effects of network neutrality when it is defined as a zero-price rule. They directly model a service provider as a two-sided platform with cross-group membership externalities (on the lines of Armstrong, 2006) that sets prices for consumers as well as providers. They relate the welfare effects of a zero-price rule to the relative size of cross-group externalities and the product differentiation parameters.

The outline of the paper is as follows. We describe the baseline model with consumers facing two-part tariffs from the monopoly service provider in Section 2. A complete information welfare analysis and the equilibrium analysis follow in Sections 2.1 and 2.2. Welfare properties of the equilibrium in the baseline model are developed in Section 2.3. The model is generalized to nonlinear pricing in Section 3, equilibrium analysis follows in Section 3.1. The welfare properties of the equilibrium are analyzed in Section 3.2. We end with suggesting some possible extensions or variants of the model that could yield further insights.

2. Consumers facing two-part tariffs from service provider

There is a single good called content[7] available for consumption in continuous quantities. There are a large but finite number of potential content providers and a continuum of heterogenous consumers indexed by a taste parameter $t \in (0, \infty)$. All content providers are identical. There is a distribution $F(t)$ of tastes among the population of consumers with the associated density $f(t)$. There is a monopoly service provider that provides last mile broadband connection to consumers and allows content providers to transmit their data through its communication channels that have a certain capacity K. The capacity is the maximum throughput (maximum rate of data transfer) of the service provider's communication channels. It is understood that any given content possibly traverses many networks before reaching the consumer through the last-mile provider. Fig. 1 shows the salient features of the model in a compact representation.

Consumers demand broadband connection but only as an input to consume content on the internet. This is the complementarity feature. The market for content is modeled as a homogenous goods market with all content providers supplying similar content. This is, of course, a modeling assumption; but one that yields rich insights as later sections will demonstrate. Let the usage price in the content market be s. Consumers are modeled as making usage decisions given the prices they face. Let $x(t)$ denote the volume of content that consumer of type t consumes. The aggregate consumer usage X determines the utilization rate of the service provider's communication channels. The congestion delay d is directly proportional

[7] Consumers are thought of as consuming data packets rather than, say, consuming Netflix content or Amazon Prime content. However, cloud storage platforms have quantity based pricing and they also happen to be a good example of content providers with homogenous goods.

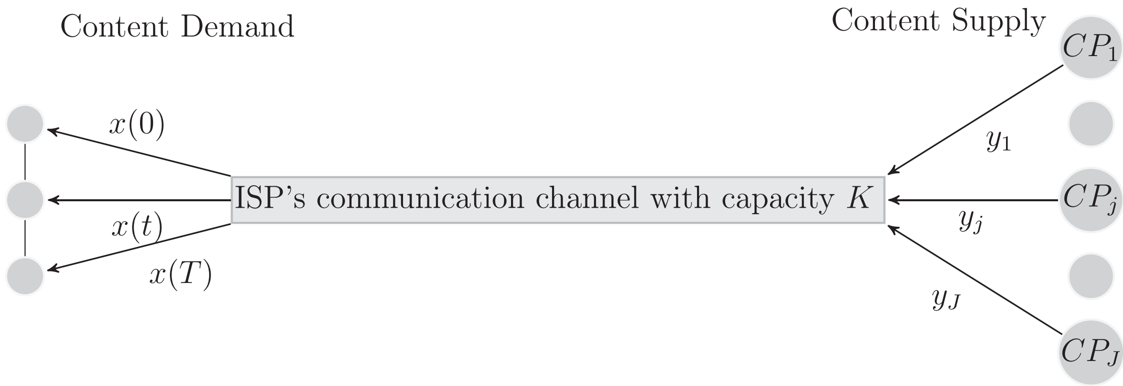

Fig. 2. An illustration of content flows in the model.

to the utilization rate which is defined as a simple ratio of total usage over capacity X/K and setting the congestion delay d equal to the utilization rate, that is, $d = X/K$. Consumers' utilities depend on the volume of content that they consume and the congestion delay that they experience. We specify consumer of type t's utility as quasilinear in money, and congestion costs as lumpsum costs.

We borrow the specification of quasilinear preferences and modeling of congestion costs from MacKie-Mason and Varian (1995) which provides a simple microfoundation for consumer demand in our model. The utility of type t consumer is given by

$$U(x, d, P_1, p_1, s; t) = \frac{1}{t}u(x) - c(d) - P_1 - p_1 x - sx, \tag{1}$$

where (P_1, p_1) is the two-part tariff[8] charged by the service provider to the consumers wherein P_1 is the access price and p_1 is the usage price, s is the usage price that a consumer pays to the content providers, $\frac{1}{t}u(x)$ is the value that consumer of type t derives from consuming a volume x of content and $c(d)$ is the cost of experiencing a congestion delay of d. Assume $u(.)$ is differentiable, strictly increasing and strictly concave; $c(.)$ is differentiable, strictly increasing and strictly convex, $u(0) = c(0) = 0$; $\lim_{x\to 0} u'(x) = \infty$ and $\lim_{x\to\infty} u'(x) = 0$.

Let us now describe the supply side of the model. Shapiro and Varian (1998) provide a suggestion for modeling content providers' costs. As purveyors of information good, their cost structure is well modeled by high fixed costs and low marginal costs. We follow their suggestion by assuming a fixed cost G and a marginal cost of 0 for content providers. The revenue model of content providers also varies in reality. Some of them rely on direct subscription prices to consumers, others depend on advertising revenues while some depend on a combination of both. We model content providers as deriving their revenues from a direct usage price s to consumers. This is of course a modeling assumption but not an unreasonable one.[9] The usage price is endogenously determined by the market structure. Let y_j be the volume of content supplied by provider j and p_2 be the linear termination fee charged by the service provider for access to its customers. If J providers are active in the market, then for every $j = 1, \ldots, J$, we write content provider j's profit as

$$\pi_C^j(P_1, p_1, p_2, (y_j)_{j\in J}) = (s - p_2)y_j - G; \quad G > 0 \tag{2}$$

Modeling the market structure for content providers is a challenge given the wide variety of content we observe on the internet. In this model, we abstract from all such heterogeneity and simply view the content providers as competing suppliers of information packets or data. Their interaction is modeled as the classic Cournot supply game[10] We also assume, as in the Cournot model, that the price s of content adjusts to clear the content market. The hypothesis of market clearing needs some explanation. We think of content providers as having created a stock of content after expending a fixed cost. The market to which we apply the market clearing condition is that of content flow or data flow. So it is useful to think of demand flows coming from consumers and supply flows coming from content providers. Fig. 2 provides an illustration of content flows in the model. The market clearing (in flows) condition is given by Eq. (3) where T is the marginal consumer who connects to the service provider.

$$X(p_1, s, T) = \int_{t=0}^{T} x(t) f(t)\, \mathrm{d}t = \sum_{j=1}^{J} y_j = Y(s, p_2) \tag{3}$$

[8] For instance, mobile data plans are often involve a fixed rental price and a usage based price.

[9] With histories of users' experience and online rating systems working to reduce informational asymmetries and with widespread penetration of credit cards and electronic payment systems even in developing countries, the arguments of Lee and Wu (2009), pp. 65–66 citing these frictions on payments between providers and consumers do not look very convincing. This is not to detract from their argument that a lot of content is free. However, much of it comes at a price.

[10] In Appendix 1, we model the content provider market alternatively as a differentiated product Bertrand duopoly.

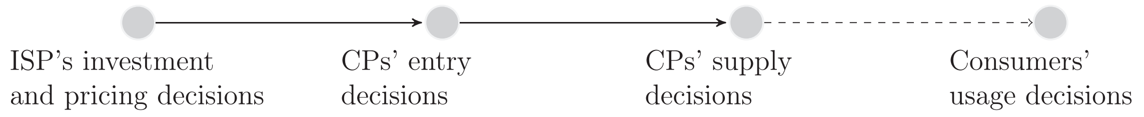

Fig. 3. Timeline of decisions in the model. The acronyms ISP and CP stand for internet service provider and content provider respectively. Consumers are modeled as nonstrategic and so theirs is not a decision node in a strict sense.

The introduction of two-part tariffs makes the consumer market size endogenous to the model. On the content provider side, we augment the Cournot supply stage with an entry stage at which potentially large number of providers make their entry decisions. This makes the provider market size endogenous as well. In other words, this makes the model symmetric[11] with respect to endogeneity of market size on the two sides. Moreover, it is natural to be interested in how the service provider's decisions shape the competitive entry decisions of content providers.

Let $g(K, X)$ denote the service provider's cost of supporting an aggregate usage X through an installed capacity K. We assume that $g(K, X) = g_0(K) + g_1 X$ where $g_0(K)$ is the fixed cost of installed capacity K and $g_1 > 0$ is the constant marginal cost of aggregate usage. The service provider's profit when T is the marginal consumer is given by

$$\begin{aligned}\pi_I(P_1, p_1, p_2, (y_j)_{j\in J}) &= P_1 F(T) + p_1 X + p_2 \sum_{j\in J} y_j - g(K, X) \\ &= P_1 F(T) + p_1 X + p_2 Y - g_0(K) - g_1 X \qquad (4)\end{aligned}$$

We posit a multi-stage game formulation which is schematically shown in Fig. 3. In the first stage, the service provider moves to choose a level of capacity as well as set prices for the consumers and for the content providers. The second stage is the entry stage where a large number of content providers move simultaneously choosing whether to enter or not. In the third stage, the content providers who chose to enter in the second stage move simultaneously to choose how much content to supply. The continuum of consumers assumption implies that we model them as nonstrategic. The players in this game are - the service provider whose strategy choices are investment and pricing decisions (on both consumer and provider side) and whose payoff function is given by Eq. (4); and the content providers indexed by j whose strategy choice is a quantity $y_j \in \mathbb{R}$ and whose payoff function is given by Eq. (2). We will study subgame perfect equilibria of this model.

The building blocks of the model that we lay out are classic economic models themselves. The content supply subgame between content providers is modeled as the classic quantity competition model of Cournot (1838) that is played in the output market. The entire game seen at once is a sequential twist on the well known pricing game of Cournot (1838) for complementary goods. In our model, the complementary goods are the broadband connection and content. We set up the service provider as a price leader who chooses two part tariffs (P_1, p_1) on the consumer side and a usage price p_2 on the content providers side. The content providers follow up but choose prices implicitly by choosing quantities explicitly in a Cournot quantity game.

Consumer demand

Consumer of type t faces the following problem of choosing content usage $x(t)$ given the prices

$$\begin{aligned}&\max_x \quad \frac{1}{t}u(x) - c\left(\frac{X}{K}\right) - P_1 - (p_1 + s)x \\ &\text{subject to} \quad \frac{1}{t}u(x) - c\left(\frac{X}{K}\right) - P_1 - (p_1 + s)x \geq 0 \qquad (5) \\ &\qquad\qquad x \geq 0\end{aligned}$$

A typical consumer is so small that he cannot influence aggregate usage of the network. This is, of course, a consequence of the continuum assumption. So in making his own usage decisions, he simply looks at the marginal price of usage and ignores the congestion externalities he imposes on the system. Ignoring the constraints and solving type t consumer's problem yields the first order condition

$$\frac{1}{t}u'(x) = p_1 + s \qquad (6)$$

By the concavity assumption on $u(.)$, it follows that a consumer with a higher value of t consumes lesser amount of content. We can write $x(t) = (u')^{-1}(t(p_1 + s))$. The indirect utility of a consumer of type t that is directly affected by his demand is given by $v(p_1, s, t) = \frac{1}{t}u(x(t)) - (p_1 + s)x(t)$. Suppose tentatively and without justification that all consumers of types $(0, T]$ connect to the service provider such that type T is the marginal consumer. Then the aggregate demand for content is given by

$$X(p_1, s, T) = \mathbb{E}[x(t) | t \in (0, T)] = \int_0^T x(p_1 + s, t) f(t)\, dt \qquad (7)$$

[11] Nevertheless, there is a modeling asymmetry that is still present – consumers are heterogenous in tastes while content providers are identical in all respects. This may be relaxed at the cost of greater complexity.

The consumer demand exhibits a complementarity relationship between broadband connection and content. The aggregate demand $X(p_1, s, T)$ depends on p_1 and s only though the sum $p_1 + s$. From Eq. (7) and the First Fundamental Theorem of Calculus, we get $x(p_1 + s, T)f(T) = \frac{\partial X(p_1+s,T)}{\partial T}$. Substituting this in the first order condition (6) of the marginal consumer's problem gives

$$s = \frac{1}{T}u'\left(\frac{1}{f(T)}\frac{\partial X}{\partial T}\right) - p_1 \tag{8}$$

Given the concavity assumption on $u(.)$, the aggregate inverse demand curve $s(X, T, p_1)$ is given by Eq. (8) and is a standard downward sloping curve with respect to X.

2.1. Welfare optimum under complete information

In this subsection, we carry out a welfare analysis of the model under the assumption that the planner has no control over the number of content providers that enter the market. The idea is to identify the tradeoffs that a social planner would face in this model and to identify the *Pigouvian* tax which would help consumers internalize the congestion externalities. As we will find later, *Pigouvian* tax shows up in the equilibrium analysis as well. The welfare analysis has aspects of MacKie-Mason and Varian (1995) but assuming the presence of a continuum and heterogeneity of consumers and marginal service costs.

A utilitarian planner with complete information about tastes among the consumers would choose the capacity K, the marginal consumer T and a pattern of content usage $x(.)$ to maximize the sum total of payoffs of all consumers, all content providers and the service provider. This problem can be written as

$$\max_{K,T,x(.)} \quad W(K, T, x(.)) = \int_0^T \frac{1}{t}u(x(t))f(t)\,\mathrm{d}t - F(T)c\left(\frac{X(T)}{K}\right) - g_1 X(T) - g_0(K)$$
$$\text{subject to } X(T) = \int_0^T x(t)f(t)\,\mathrm{d}t$$
$$\forall t \in (0, T], \quad x(t) \geq 0$$

We assume that the integral in the objective function is well defined. As the objective function indicates, the determinants of welfare are aggregate consumer value of content, aggregate congestion cost borne by consumers, service cost of supporting the aggregate usage and the investment cost in capacity. The planner would choose $(K, T, x(.))$ that resolves the following tradeoffs.

Consumer market size T: The social value of increasing T is the marginal consumer's value of content. The social cost is a marginal increase in service expenditure as well as an increase in congestion cost that can be decoupled into a direct effect and an indirect effect. The direct effect is the congestion cost borne by the new marginal consumer while the indirect effect is the marginal increase in congestion costs borne by the inframarginal consumers due to the additional usage of the new marginal consumer. The planner's optimal choice of T balances these social costs and benefits and is given by

$$\frac{1}{T}u(x(T)) = c\left(\frac{X(T)}{K}\right) + F(T)\frac{1}{K}c'\left(\frac{X(T)}{K}\right)x(T) + g_1 x(T) \tag{9}$$

Content usage pattern $x(.)$: The social value of increasing the usage of type-t consumer is given by this consumer's marginal value $\frac{1}{t}u'(x(t))$. The social cost is given by a marginal increase in aggregate congestion costs plus marginal service costs of the internet provider. The first order condition corresponding to usage of type-t consumer may be simplified to

$$\frac{1}{t}u'(x(t)) = F(T)\frac{1}{K}c'\left(\frac{X(T)}{K}\right) + g_1 \tag{10}$$

The first term on the right hand side in Eq. (10) may also be written as $\frac{\partial c}{\partial X}F(T)$ and is the marginal social cost of congestion. This is independent of the taste of a consumer. The planner may implement the welfare-optimal pattern of usage by setting a *Pigouvian* tax given by

$$e = p_c + g_1, \tag{11}$$

where $p_c = \frac{\partial c}{\partial X}F(T) = F(T)\frac{1}{K}c'(\frac{X(T)}{K})$ may be interpreted as the congestion charge and is set to be equal to the marginal social cost of congestion. Faced with a Pigouvian tax e on usage, the decentralized solution to the consumer's problem $\max_{x(t)} \frac{1}{t}u(x(t)) - ex(t)$ is the same as the centralized solution to the social planner's problem of allocating a pattern of usage. In other words, the tax forces the consumers to internalize the congestion externalities they generate by their usage.

Capacity choice K: Using the idea that the social planner will set the congestion charge p_c to be equal to the marginal social cost of congestion, we can write the first order condition for optimal K in a more interpretable form as

$$\frac{\partial W}{\partial K} = p_c\frac{X(T)}{K} - g_0'(K) = 0 \tag{12}$$

Increasing capacity will increase welfare if and only if revenues from congestion charge make up for the cost of capacity provision (computed by valuing capacity by its marginal cost). Thus the congestion charge sends the right economic signal to the planner to expand capacity.

2.2. Equilibrium analysis

Cournot subgame for content supply

A fundamental postulate of the Cournot production game is market clearing. Using Eq. (3) in Eq. (8) gives the residual inverse demand curve for content as

$$s(Y,T,p_1) := \frac{1}{T}u'\left(\frac{1}{f(T)}\frac{\partial Y}{\partial T}\right) - p_1 =: \phi(Y,T) - p_1 \tag{13}$$

Eq. (13) which is the mathematical expression of the economic relationship of complementarity implies that the markup in this Cournot subgame depends only on the usage price level $p = p_1 + p_2$. As shown in Eq. (14), for a given value of p, the markup is invariant to the precise marginal cost (termination fee) p_2. This is unlike the traditional Cournot model and is solely driven by the complementarity.

$$s - p_2 = \phi(Y,T) - p \tag{14}$$

All content providers are alike. So given the service provider's choice of prices (P_1, p_1, p_2), a typical content provider j's profit in Eq. (2) can be rewritten as $\pi_C^j(P_1, p_1, p_2, y_j, Y) = (s(Y,T,p_1) - p_2)y_j - G$. Ignoring the participation constraints and using Eq. (14), the aggregate supply of content can be deduced from the corresponding first order conditions[12] as

$$Y\frac{\partial \phi}{\partial Y} = J(p - \phi(Y,T)) \tag{16}$$

So the aggregate content supply $Y(p,T,J)$ is determined by Eq. (16) if $p_2 < s$ and is equal to zero if $p_2 \geq s$. This is because when the service provider chooses a termination fee $p_2 > s$, it violates the participation constraints of the content providers as they face revenue losses. The equilibrium in this Cournot subgame[13] is symmetric where every provider supplies the same amount of content. Let $y(J)$ be the equilibrium content supply by every provider. The aggregate content supply when there are J firms is then $Y = Jy(J)$. Let $\pi_C(J)$ be the equilibrium profit of a content provider.

Entry subgame

When there are no barriers to entry other than the fixed costs G, the free entry equilibrium will have the number of firms that is given by the zero profit condition (ignoring the integer constraint on the number of firms).

$$\pi_C(J) = (s - p_2)\frac{Y}{J} - G = 0 \tag{17}$$

Eqs. (14) and (17) imply that given the service provider's strategy choice of (K, P_1, p_1, p_2), the equilibrium entry is determined from the following equation in J

$$J = \frac{(\phi(Y) - p)Y}{G} \tag{18}$$

where $Y(p,T,J)$ is written as Y with its arguments suppressed and is determined in the Cournot subgame by Eq. (16). The first things to note is that the markup given by Eq. (14) must be positive to make entry feasible. Next, the equilibrium entry also depends only on the usage price level and not the price structure. The aggregate supply Y of content in a free entry equilibrium is determined by substituting Eq. (18) in Eq. (16) and is determined from[14]

$$\frac{\partial \phi(Y)}{\partial Y} = -\frac{(\phi(Y) - p)^2}{G} \tag{20}$$

Eq. (20) makes it clear that for a given usage price level p, the aggregate content flow is insensitive to the way p is split up for both sides.

Service provider's capacity choice and price leadership problem

The service provider takes into account the supply outcomes of the entry augmented Cournot subgame. It also takes into account that the price s will adjust to clear the content market. When the aggregate demand X is specified by Eq. (8) and

[12] These conditions are that

$$\forall j = 1, \ldots, J, \quad y_j\frac{\partial s}{\partial Y} = p_2 - s(Y,T,p_1) \tag{15}$$

[13] In this symmetric model, our modeling assumption of zero variable costs for content provider firms coupled with a downward sloping and continuous inverse demand implies equilibrium existence in the Cournot subgame by the results of McManus (1962), McManus (1964) and Roberts and Sonnenschein (1976).

[14] It may be of interest to look at the comparative statics of content price with respect to the fixed cost of content industry. This is easily carried out by rewriting (20) using (14) as:

$$(s - p_2)^2 = G\left(-\frac{\partial \phi}{\partial Y}\right) \tag{19}$$

The sign of $\frac{\partial \phi}{\partial Y}$ is the same as the sign of u'' which is negative due to the concavity assumption on u. It follows, therefore, that s varies positively with G.

the aggregate supply Y is specified by Eq. (20), the service provider's problem in a subgame perfect equilibrium is

$$\max_{K,P_1,p_1,p_2} \quad P_1F(T)+p_1X(p_1,s,T)+p_2Y(p_2,s)-g_0(K)-g_1Y$$

subject to Market Clearing as specified by Eq. (3)

Participation Constraints for all consumers as specified by Eq. (5)

This formulation of the problem makes it transparent that the service provider's ability to influence content demand is complicated by the complementarity relationship between the broadband connection and content. In addition, it must respect the participation constraints of consumers in equilibrium. Moreover, its ability to directly reduce the margin of the content providers and thereby affect the content supply is constrained by the need to ensure participation constraints are met for the content providers which is already accounted for in the equilibrium content supply $Y(p_2, s)$.

Now from the participation constraint (5) for the marginal consumer's problem, it is clear that the service provider will set an access price P_1 so that

$$v(p_1,s,T)=c\left(\frac{X(p_1,s,T)}{K}\right)+P_1 \tag{21}$$

The existence and uniqueness of a marginal type T such that the above equation holds is guaranteed because $v(p_1,s,T)$ is a decreasing function of T with $\lim_{T\to 0} v(p_1,s,T)=\infty$ while $c(\frac{X(p_1,s,T)}{K})$ is an increasing function of T with $\lim_{T\to 0} c(\frac{X(p_1,s,T)}{K})=0$. All consumers of type $t\leq T$ connect because for any such type $v(p_1,s,t)\geq v(p_1,s,T)$.

Using Eq. (21) to substitute for optimal access price P_1 and the market clearing postulate $X=Y$, we may rewrite the service provider's problem in a equilibrium as that of choosing the capacity K, the marginal consumer T and the total usage price p

$$\max_{K,T,p} \quad \left[v(p,T)-c\left(\frac{Y(p,T)}{K}\right)\right]F(T)+(p-g_1)Y(p,T)-g_0(K)$$

This version of the service provider's problem makes it clear that its profit in equilibrium depends only on usage price level p and not on the precise usage price structure $(p_1, p_2): p_1+p_2=p$; and is similar to that of a monopolist choosing the optimal two-part tariff as exposited in Varian (1989) (Oi, 1971 and Schmalensee (1981) are early references on two-part tariffs). However, there are two principal differences. First, while increasing the aggregate content flow increases the consumer surplus, it also increases the congestion costs, thereby reducing the ability to extract that surplus owing to the need to respect the participation constraints of consumers. Second, there is a limit on the usage price it can set for content providers owing to the need to respect their participation constraints.

A triple (K, T, p) in an equilibrium must satisfy the necessary first order conditions for optimality. The optimality conditions with respect to T and p look similar to those derived in the context for optimal two-part tariff in Varian (1989) except in two respects. One, the presence of congestion externalities acts to reduce the service provider's profit margin by reducing the gross consumer surplus. Two, the optimal (K, T, p) must respect content providers' participations constraints.

Capacity choice K: The optimality condition with respect to K is exactly the same as the optimality condition (12) for the social planner in a welfare optimum. The monopolist's incentives to invest in capacity is exactly the same as the social planner's incentives to invest in capacity.

Consumer market size T: The optimality condition[15] with respect to T can be intuitively understood by rewriting it as (23) in terms of the access price using Eq. (21) which says $v-c=P_1$.

$$\frac{\partial P_1}{\partial T}F(T)+P_1f(T)+(p-g_1)\frac{\partial Y}{\partial T}=0 \tag{23}$$

The monopolist faces a tradeoff. Increasing the consumer market size benefits him because he can make more money from both usage and access prices. However, the additional consumer's value of content is less than the existing consumers. Moreover, it increases congestion on the network. This reduces the access price that the monopolist can charge.

Usage price level p: using Roy's identity[16]; Chain Rules $\frac{\partial c}{\partial p}=\frac{\partial c}{\partial Y}\frac{\partial Y}{\partial p}$ and $\frac{\partial c}{\partial T}=\frac{\partial c}{\partial Y}\frac{\partial Y}{\partial T}$; definitions $\eta(Y,p)=-\frac{pY'(p)}{Y(p)}$ of the elasticity of content flow with respect to the total usage price, and of Pigouvian tax e from Eq. (11), the optimality condi-

[15] The first order condition is

$$\frac{\partial \pi_I}{\partial T}=\left[\frac{\partial v(p,T)}{\partial T}-\frac{\partial c(Y(p,T)/K)}{\partial T}\right]F(T)+[v(p,T)-c(p,T)]f(T)+(p-g_1)\frac{\partial Y}{\partial T}=0 \tag{22}$$

[16] Consumers in our model have quasilinear utility and therefore for given price p and income m, an indirect utility function of the form $V(p,m,t)=v(p,t)+m$. By Roy's Identity, we have $x(p,t)=-\frac{\partial v(p,t)}{\partial p}$.

tion[17] with respect to p reads as

$$\frac{p-e}{p} = \frac{1}{\eta(Y,p)}\left(1-\frac{x(p,T)}{X(p,T)/F(T)}\right) \tag{25}$$

Eq. (25) resembles the first-order condition for optimal usage price as developed by Varian (1989) in the classic setting of a monopoly but differs from it in two crucial respects. One, unlike the classic monopoly setting, our model has two 'sides' and it is the usage price level and not the price on any one side that obeys the Lerner's formula. Two, the markup in (25) is over the Pigouvian tax e (which is precisely the marginal social cost of usage $p_c + g_1$ from Eq. (11)) that a social planner would set rather than over the private marginal cost of the monopolist. The contrast can be seen most clearly when the marginal consumer's demand is the same as the average consumer which happens when all consumers are identical. In this case, as per the classic usage pricing formula exposited in Varian (1989), the monopolist sets a usage price equal to the marginal cost of usage that it bears. In our model, it sets a usage price equal to the Pigouvian tax. We have derived the following result about equilibrium in the model.

Theorem 1. *With consumers facing two-part tariffs and content providers facing a linear termination fee, we have*

(i) *the monopolist service provider's profits are a function only of the usage price level $p := p_1 + p_2$ on the network. For a given usage price level p, the profits are invariant to the usage price structure $\{(p_1, p_2) : p_1 + p_2 = p\}$.*
(ii) *the usage price level p must satisfy Eq. (25), a Lerner-style formula in which the markup is over the Pigouvian tax.*
(iii) *the choice of an access price P_1 is equivalent to the choice of a marginal consumer T which delimits the consumer market size and must satisfy Eq. (22) in equilibrium.*

Theorem 1 enables us to deduce the consequences of a neutrality regulation that takes the form of zero-price rule $p_2 = 0$ in the model. This constrains the service provider to one-sided pricing and does not require any information for an effective implementation. We have the following corollary.

corollary 1.1 For any equilibrium E_1 in which the service provider earns profit π_I with a usage price level p that features a positive termination fee, there exists another equilibrium E_2 in which the service provider earns the same profit π_I with the same usage price level p but with zero termination fee. In other words, a zero-price rule is neutral in this model.

2.3. Welfare properties of equilibrium

Welfare is given by consumers' surplus plus the service provider's profits plus the aggregate profits of the content providers:

$$W(K,T,p) = \int_0^T v(p,t)f(t)\,\mathrm{d}t - c\left(\frac{Y}{K}\right)F(T) + (p-g_1)Y(p,T) - g_0(K) + (s-p_2)Y(p,T) - JF \tag{26}$$

At the free-entry symmetric sequential equilibrium, $Y = Jy(J)$ and $\pi_C(J) = 0$. So Eq. (26) may be rewritten as

$$\begin{aligned} W(K,T,p) &= \int_0^T v(p,t)f(t)\,\mathrm{d}t - c\left(\frac{Y}{K}\right)F(T) + (p-g_1)Y(p,T) - g_0(K) + J\pi_C(J) \\ &= \int_0^T v(p,t)f(t)\,\mathrm{d}t - c\left(\frac{Y}{K}\right)F(T) + (p-g_1)Y(p,T) - g_0(K) \end{aligned} \tag{27}$$

Since profits of every content provider is driven to zero in a free-entry sequential equilibrium, those profits do not figure in the social planner's objective at equilibrium. So equilibrium welfare in this model is the consumers' surplus plus the monopolist service provider's profits. The situation is much like the classic setting of an ordinary monopolist pricing consumers via two-part tariffs as exposited in Varian (1989). The difference, however, is that the usage price p includes the termination fee p_2 on the provider side. A change in p_2 affects the flow of content on the network, thereby affecting both the service provider's profits and the consumers' surplus. Differentiating with respect to K, T and p and using Roy's identity again, we have

$$\frac{\partial W(K,T,p)}{\partial K} = \frac{F(T)}{K}c'\left(\frac{Y}{K}\right)\frac{Y}{K} - g_0'(K) \tag{28}$$

$$\frac{\partial W(K,T,p)}{\partial T} = \left[v(p,T) - c\left(\frac{Y}{K}\right) + (p-g_1)x(p,T)\right]f(T) \tag{29}$$

[17] The optimality condition for p is

$$\frac{\partial \pi_I}{\partial p} = \left[\frac{\partial v(p,T)}{\partial p} - \frac{\partial c(Y(p,T)/K)}{\partial p}\right]F(T) + (p-g_1)\frac{\partial Y}{\partial p} + Y(p,T) = 0 \tag{24}$$

$$\begin{aligned}\frac{\partial W(K,T,p)}{\partial p} &= -\int_0^T x(p,t)f(t)\,dt - \frac{F(T)}{K}c'\left(\frac{Y}{K}\right)\frac{\partial Y}{\partial p} + (p-g_1)\frac{\partial Y}{\partial p} + Y(p,T)\\ &= \frac{\partial Y}{\partial p}\left(p - g_1 - \frac{F(T)}{K}c'\left(\frac{Y}{K}\right)\right)\end{aligned} \tag{30}$$

The last equation follows due to market clearing. Evaluating these partial derivatives at (K^e, T^e, p^e), the equilibrium values of (K, T, p) that are given by Eqs. (12), (22) and (24), we have

$$\frac{\partial W(K^e,T^e,p^e)}{\partial K} = 0 \tag{31}$$

$$\frac{\partial W(K^e,T^e,p^e)}{\partial T} = \left[\frac{\partial c(p^e,T^e)}{\partial T} - \frac{\partial v(p^e,T^e)}{\partial T}\right]F(T^e) > 0 \tag{32}$$

$$\frac{\partial W(K^e,T^e,p^e)}{\partial p} = x(p^e,T^e)F(T^e) - X(p^e,T^e) \tag{33}$$

From Eq. (31), we conclude that the equilibrium provision of capacity by the service provider is socially optimal. From Eq. (32), the equilibrium consumer market size is too small relative to the welfare optimal level because $\frac{\partial c(p,T)}{\partial T} > 0$, $\frac{\partial v(p,T)}{\partial T} < 0$ and $F(T) > 0$. Moreover, from Eq. (33), we have

$$\frac{\partial W(K^e,T^e,p^e)}{\partial p} > 0 \quad \text{if and only if} \quad x(p^e,T^e) - \frac{X(p^e,T^e)}{F(T^e)} > 0 \tag{34}$$

This analysis gives price regulation, a potentially welfare-enhancing role. We have the following summary result about the welfare properties of equilibrium.

Theorem 2. *With consumers facing two part tariffs and content providers facing linear usage pricing, the monopolist service provider's*

(i) *equilibrium provision of capacity is socially optimal.*
(ii) *choice of consumer market size in equilibrium is too small relative to the social optimum.*
(iii) *equilibrium usage price level is too low relative to the socially optimal level if and only if the marginal consumer demand is bigger than the average consumer demand.*

We close this section with an example that is a special case of the model in which the service provider's optimal usage price level is determined by the classic Lerner's formula for a monopolist.

Example: Identical consumers facing linear usage pricing from service provider. Suppose there are a unit mass of consumers with identical tastes and the service provider charges a linear usage price but no access price to consumers. With a continuum of consumers, the individual usage x is also the aggregate usage X. A typical consumer's utility is given by $u(x) - c(d) - p_1 x - sx$ where symbols inherit their earlier meaning. The service provider's profit is given by $\pi_I(p_1, p_2, (y_j)_{j\in J}) = p_1 X + p_2 \sum_{j\in J} y_j - g(K, X)$. The rest of the model is the same. In this symmetric model, issues of consumer market access do not arise. Universal service to consumers is ensured as long as

$$u(X) - c\left(\frac{X}{K}\right) - (p_1 + s)X \geq 0 \tag{35}$$

Assuming universal service provision by the service provider, the aggregate (or individual) demand $X(s, p_1)$ is determined from the consumer's problem as

$$p_1 + s = u'(X) \tag{36}$$

The aggregate supply is determined in the entry-augmented Cournot game among the content providers and by an analysis that closely parallels the corresponding analysis in Section 2.2, is given by

$$u''(Y) = -\frac{(u'(Y) - p)^2}{G} \tag{37}$$

when $p_2 \leq s$ and equals 0 otherwise. Like in the baseline model, for a given usage price level p, the aggregate content supply is insensitive to the price structure (p_1, p_2). Consequently, the resulting market is not a two-sided market in the sense of Rochet and Tirole (2006). When the aggregate demand is specified by Eq. (36) and aggregate supply is specified by Eq. (37), the service provider's problem in a subgame perfect equilibrium is

$$\begin{aligned}&\max_{K,p_1,p_2} && p_1 X(s,p_1) + p_2 Y(s,p_2) - g_0(K) - g_1 Y\\ &\text{subject to} && \text{Market Clearing as specified by Eq. (3)}\\ & && \text{Universal Service Constraint as specified by Eq. (35)}\end{aligned}$$

We can rewrite the service provider's problem in an equilibrium as that of choosing the price level p.

$$\max_{K,p} \quad (p-g_1)Y(p)-g_0(K)$$
$$\text{subject to} \quad \text{Universal Service Constraint}$$

Ignoring the constraint, this is the standard monopoly problem and the optimal price level is determined by the standard markup formula

$$\frac{p-g_1}{p}=\frac{1}{\eta} \tag{38}$$

where $\eta=-\frac{pY'(p)}{Y(p)}$ is the elasticity of content flow with respect to the usage price level. The monopolist then simply chooses the capacity K so as to satisfy the universal service constraint given by Eq. (35).

3. Consumers facing nonlinear pricing from service provider

In this section, we generalize the baseline model of Section 2 to the case where the service provider charges a nonlinear price $\mathfrak{p}_1(.)$ to consumers. As before, the type of a consumer is his taste parameter t. A higher t reflects both a lower absolute value of content as well as a lower marginal value of content. The service provider knows $u(.)$ and $c(.)$ but does not know the taste parameter t. It believes t is distributed according to the distribution function F with the associated density f over the positive real line. The utility of a consumer of taste parameter t when he consumes $x(t)$ amount of content and experiences a delay of d is given by

$$U(t)=\frac{1}{t}u(x(t))-c(d)-\mathfrak{p}_1(t)-sx(t) \tag{39}$$

where symbols inherit the meaning they held in Section 2.

The content providers side is still modeled as an entry-augmented symmetric Cournot oligopoly. When J is the equilibrium number of firms that enter, the market clearing condition is

$$X(\mathfrak{p}_1(.),s,T):=\int_0^T x(t)f(t)\,\mathrm{d}t=\sum_{j=1}^{J}y_j=:Y(s,p_2) \tag{40}$$

The service provider's profit when T is the marginal consumer is given by

$$\pi_I(\mathfrak{p}_1(.),p_2,(y_j)_{j\in J})=\int_0^T \mathfrak{p}_1(t)f(t)\,\mathrm{d}t+p_2\sum_{j\in J}y_j-g(K,X) \tag{41}$$

The welfare analysis of this model is the same as in Section 2.1 and is omitted here.

The derivation of consumer demand in this model makes use of the classic theory of screening (Mussa and Rosen, 1978; Maskin and Riley, 1984). A direct mechanism is a pair comprising an allocation rule $x:(0,T]\mapsto\mathbb{R}_+$ and a pricing rule $\mathfrak{p}_1:(0,T]\mapsto\mathbb{R}$. The derivation of the pricing rule in terms of the allocation rule is quite standard in the theory of screening with one-dimensional types. In our model however, nonlinear pricing is only one aspect and it is intertwined with complementarity and congestion externalities. Nevertheless, a derivation of the inverse demand curve (42) is provided in Appendix 2.

$$s=\frac{1}{T}u'(x(T))-\frac{\partial\mathfrak{p}_1(T)}{\partial x(T)} \tag{42}$$

This inverse demand curve is consistent with the baseline model featuring two-part tariffs. In that case, $\mathfrak{p}_1(T)=P_1+p_1x(T)$; so Eq. (42) reduces to Eq. (8). As in the baseline model, for a given value of $\mathfrak{p}_1(T)$, the inverse demand curve is a downward sloping curve because of the concavity of $u(.)$. Note that s depends on the pricing rule $\mathfrak{p}_1(.)$ only through the price $\mathfrak{p}_1(T)$ set for the marginal consumer.

3.1. Equilibrium analysis

Supply subgame

In the entry-augmented Cournot subgame for content supply, the residual inverse demand curve for content is obtained from Eq. (42) and the market clearing condition as[18]

$$s(Y,T,\mathfrak{p}_1(T))=\frac{1}{T}u'\left(\frac{1}{f(T)}\frac{\partial Y}{\partial T}\right)-\frac{\partial\mathfrak{p}_1(x(T)=\frac{1}{f(T)}\frac{\partial Y}{\partial T})}{\partial x(T)} \tag{43}$$

[18] This involves an application of the First Fundamental Theorem of Calculus to the definition of aggregate demand.

The analysis remains the same as the baseline model so that the aggregate supply $Y(\mathfrak{p}_1(T), p_2, T)$ of content in equilibrium is determined from Eq. (44).

$$\frac{\partial s(Y)}{\partial Y} = -\frac{(s(Y) - p_2)^2}{G} \tag{44}$$

The important difference here is that unlike the baseline model, the marginal price of usage on the consumer side is not constant. Consequently, the crucial result in the baseline model – for a given usage price level, aggregate content flow is insensitive to the precise split on both sides – no longer obtains here.

Service provider's capacity choice and price leadership problem

When the aggregate demand is specified by Eq. (42) and aggregate supply is specified by Eq. (44), the service provider's problem in a subgame perfect equilibrium is

$$\begin{aligned} [OPT1] \quad & \max_{K,T,\mathfrak{p}_1(.),p_2} \quad \int_0^T \mathfrak{p}_1(t) f(t)\, \mathrm{d}t + (p_2 - g_1) Y(\mathfrak{p}_1(T), p_2, T) - g_0(K) \\ & \text{subject to} \quad (x(.), \mathfrak{p}_1(.)) \text{is incentive compatible and individually rational} \\ & \qquad\qquad\quad \text{Market Clearing as specified by Eq. (40)} \end{aligned}$$

We can transform the service provider's problem given by $[OPT1]$ in which it chooses, among other things, a price schedule $\mathfrak{p}_1(.)$ on the consumer side into $[OPT2]$ where it chooses the allocation rule $x(.)$ on the consumer side instead of a price schedule. The argument that is needed for this transformation is standard in the theory of screening and we show it in the Appendix 2 in the context of this model.

$$\begin{aligned} [OPT2] \quad & \max_{K,T,x(.),p_2} \quad \int_0^T \left[\left(\frac{1}{t} - \frac{F(t)}{t^2} \right) u(x(t)) - s x(t) - c\left(\frac{X}{K} \right) \right] f(t)\, \mathrm{d}t + (p_2 - g_1) Y(\mathfrak{p}_1(T), p_2, T) - g_0(K) \\ & \text{subject to} \quad x(.) \text{ is decreasing and Market Clearing} \end{aligned}$$

When looking to derive optimality conditions for $(K, T, x(.))$ in an equilibrium, we use the market clearing condition to write the optimization problem as

$$\begin{aligned} [OPT3] \quad & \max_{K,T,x(.)} \quad \int_0^T \left[\left(\frac{1}{t} - \frac{F(t)}{t^2} \right) u(x(t)) - s x(t) - c\left(\frac{X}{K} \right) \right] f(t)\, \mathrm{d}t + (p_2 - g_1) \int_0^T x(t) f(t)\, \mathrm{d}t - g_0(K) \\ & \text{subject to} \quad x(.) \text{ is decreasing} \end{aligned}$$

Optimal allocation rule

First ignore the monotonicity constraint on $x(.)$ and find for every t, the allocation $x(t)$ that maximizes the integrand.

$$\left(\frac{1}{t} - \frac{F(t)}{t^2} \right) u'(x(t)) = g_1 + s - p_2 \tag{45}$$

Under the constraint that the markup of content providers must be positive, the right side of Eq. (45) is positive. So a necessary condition for existence of a solution to this equation is that $\left(1 - \frac{F(t)}{t}\right) > 0$. The following assumptions ensure that a unique solution $x(t)$ exists for every t and that $x(t)$ is decreasing over $(0, T]$.

Assumption 1. $\lim_{t \to 0} f'(t) = 0$.

Assumption 2. $\left(1 - \frac{F(t)}{t}\right)$ is positive and nonincreasing in t over $(0, \bar{T}]$ for some $\bar{T} > 0$.

The left side of Eq. (45) shares the interpretation of virtual marginal benefit with screening models. The subtracted term $\frac{F(t)}{t^2} u'(x(t))$ is the information rent earned by a consumer of type t. Under Assumption 1, the information rent tends to zero for a consumer whose type t approaches zero arbitrarily closely. This is the familiar 'no distortion at the top' feature of screening models where top is understood to mean a consumer with the highest marginal valuation. Assumption 1 is satisfied by the uniform and the gamma family of distributions. One way to view this assumption is that it places constraints on the distributions through which we should model taste heterogeneity in the model if we want to preserve the 'no distortion at the top' feature. For instance, it rules out the exponential distribution.

Assumption 2 is the analogue of the regularity condition in simple screening models. Under this assumption, the left side of Eq. (45) is positive and tends to infinity when $x(t)$ tends to zero; moreover, it is decreasing in t over $(0, T]$. This implies existence and uniqueness of $x(t)$ for every t and also that $x(.)$ is decreasing. Assumption 2 is satisfied by the uniform distribution (for $\bar{T} > 1$) and many members of the gamma family of distributions.[19]

Eq. (45) can be conveniently restated in words as follows

$$\begin{aligned} &\text{Virtual Marginal Benefit of Type-}t\text{ Consumer} - \text{Content Providers' Markup} \\ &\qquad\qquad\qquad\qquad = \text{Service Provider's Marginal Costs} \end{aligned} \tag{46}$$

[19] The value of $\bar{T}$ for a gamma distribution depends on the two parameters (α, β)

The optimal allocation rule in the present model as stated in words by Eq. (46) is easily contrasted with the corresponding rule in the classic screening model which equates the virtual marginal benefit of the consumer to the marginal cost of the monopolist. In the present model, the monopolist service provider is forced to cede some economic pie not only to the consumers as informational rent but also to content providers as markup due to the complementarity relationship that exists between the two.

Optimal consumer market size

An application of Leibniz Rule in [OPT3] gives the optimality condition[20] for T. An equivalent condition that is closer to model primitives is obtained[21] as

$$\frac{\partial \mathfrak{p}_1(t=T)}{\partial T}F(T) + \mathfrak{p}_1(T)f(T) + (p_2 - g_1)\frac{\partial X}{\partial T} = 0 \tag{48}$$

Eq. (48) establishes consistency of this optimality condition with its counterpart, Eq. (23), in the baseline model with two-part tariffs. At the same time, it reveals that the tradeoffs involved here are no different than that model. The optimal market size is one that resolves the following tradeoff - on the one hand, a marginal increase in consumer market size leads to a marginal increase in revenues of the service provider on account of its ability to price the increased content flows on both sides; on the other hand, it leads to a marginal decrease in consumer prices.

Optimal capacity investment

Using the definition of p_c in Eq. (11) as the marginal social cost of congestion, the optimal capacity investment solves the first order condition

$$p_c X = g_0'(K)K \tag{49}$$

Optimal termination fee

When looking to derive optimality conditions for p_2 in an equilibrium, we use the market clearing condition to obtain the following equivalent reformulation of [$OPT2$].

$$[OPT4] \quad \max_{p_2} \quad \int_0^T \left(\frac{1}{t} - \frac{F(t)}{t^2}\right) u(x(t)) f(t)\,\mathrm{d}t - c\left(\frac{Y}{K}\right) + (p_2 - g_1 - s)Y - g_0(K)$$

The tradeoffs involved in setting an optimal termination fee are intuitive. The monopolist service provider knows that changing the termination fee affects both the supply side and the demand side. There is the temptation for making more money by charging a higher fee. This results in reduced content flows, both directly on the supply side and indirectly on the demand side by increasing content price s. This hurts the service provider because they lead to lower sales on both sides and also lowers the flexibility to increase consumer price.

With the definition of Pigouvian tax e from Eq. (11), the optimality condition that resolves these tradeoffs resembles a Lerner's formula, although a bit more complicated than in the two-part tariff model. The markup in the Lerner index is over $e + s\left(1 - \frac{1}{\eta(Y,s)}\right)$ which is the Pigouvian tax plus content price that is scaled down by a factor related to demand elasticity. The formula is

$$\frac{p_2 - e - s\left(1 - \frac{1}{\eta(Y,s)}\right)}{p_2} = \frac{1}{\eta(Y, p_2)}, \tag{50}$$

where $\eta(Y,s)$ is the elasticity of content demand with respect to content price s and $\eta(Y,p_2)$ is the elasticity of content supply with respect to the termination fee p_2. Note that the supply itself depends not only on the price p_2 charged to the provider side but also on other endogenous variables like the pricing rule $\mathfrak{p}_1(.)$ set for the consumer side and the consumer market size T. Eq. (50) says that as elasticity of content demand increases, the equilibrium markup in termination fee decreases.

Unlike the usage price level in the two-part tariff model, it is only the price on the provider side i.e. the termination fee that must obey a Lerner's formula. Again, unlike the two-part tariff model, heterogeneity in consumer taste plays no role in this Lerner's formula. In this model, it is in the determination of nonlinear pricing rule on the consumer side where consumer heterogeneity matters. Theorem 3 summarizes the salient features of the equilibrium in this model.

Theorem 3. *With consumers facing nonlinear pricing and content providers facing linear usage pricing, the monopolist's*

(i) *equilibrium nonlinear price schedule $\mathfrak{p}_1(.)$ for consumers is completely determined (up to the congestion cost) by the equilibrium usage allocation rule $x(.)$ for consumers through Eq. (75). Under Assumption 1 and 2, complementarity shapes the equilibrium usage allocation rule $x(.)$ which must satisfy Eq. (45).*

[20] The optimality condition is

$$-\frac{F(T)}{T^2}u(x(T)) + \left(\frac{1}{T}u(x(T)) - sx(T) - c\left(\frac{X}{K}\right)\right)f(T) + (p_2 - g_1)x(T)f(T) = 0 \tag{47}$$

[21] by making use of Eq. (75) in Appendix 2 to rewrite the optimality condition (47).

(ii) *equilibrium termination fee* p_2 *must satisfy Eq. (50), a Lerner-style formula in which the markup is over the Pigouvian tax plus content price that is scaled down by a factor related to demand elasticity.*

(iii) *equilibrium consumer market size must satisfy Eq. (47) and is affected both by network congestion costs and by the termination fee* p_2 *among others.*

3.2. Welfare properties of equilibrium

The equilibrium welfare analysis of this model proceeds on similar lines as in Section 2.3. Welfare is given by consumers' surplus plus the service provider's profits plus the aggregate profits of the content providers:

$$W(K,T,x(.),p_2)=\int_0^T\left(\frac{1}{t}u(x(t))-\mathfrak{p}_1(t)-sx(t)\right)f(t)\,\mathrm{d}t-c\left(\frac{Y}{K}\right)F(T)+\int_0^T\mathfrak{p}_1(t)f(t)\,\mathrm{d}t$$
$$+(p_2-g_1)Y-g_0(K)+(s-p_2)Y-JF \tag{51}$$

At the free-entry symmetric sequential equilibrium, $Y=Jy(J)$ and $\pi_C(J)=0$. Moreover, the social planner is subject to feasibility (which is market clearing in the decentralized model) as well. These observations imply the social welfare in Eq. (51) may be rewritten as

$$\begin{aligned}W(K,T,x(.),p_2)&=\int_0^T\left(\frac{1}{t}u(x(t))-sx(t)\right)f(t)\,\mathrm{d}t-c\left(\frac{Y}{K}\right)F(T)+(p_2-g_1)Y-g_0(K)+J\pi_C(J)\\&=\int_0^T\left(\frac{1}{t}u(x(t))-sx(t)\right)f(t)\,\mathrm{d}t-c\left(\frac{Y}{K}\right)F(T)+(p_2-g_1)Y-g_0(K)\\&=\int_0^T\frac{1}{t}u(x(t))f(t)\,\mathrm{d}t-c\left(\frac{Y}{K}\right)F(T)-g_1Y-g_0(K)-(s-p_2)Y\end{aligned} \tag{52}$$

We will evaluate the partial derivatives of welfare with respect to each of the variables at their equilibrium values denoted by $(K^e,T^e,x^e(.),p_2^e)$ which are given by Eqs. (49), (47), (45) and (50) to find out whether the equilibrium values are too low or too high relative to the social optimum.

The partial derivative of welfare with respect to capacity and when evaluated at the equilibrium value $(K^e,T^e,x^e(.),p_2^e)$ is given by

$$\frac{\partial W(K,T,x(.),p_2)}{\partial K}=\frac{F(T)}{K}c'\left(\frac{Y}{K}\right)\frac{Y}{K}-g_0'(K) \tag{53}$$

$$\frac{\partial W(K^e,T^e,x^e(.),p_2^e)}{\partial K}=0 \tag{54}$$

The equilibrium provision of capacity by the service provider is socially optimal. This conclusion is consistent throughout the models explored in this paper.

When finding the partial derivative of welfare with respect to consumer market size T or the allocation $x(t)$, it is convenient to view the welfare in Eq. (52) as

$$W(K,T,x(.),p_2)=\int_0^T\left(\frac{1}{t}u(x(t))+(p_2-s-g_1)x(t)-c\left(\frac{X}{K}\right)\right)f(t)\,\mathrm{d}t-g_0(K) \tag{55}$$

We can then find that

$$\frac{\partial W(K,T,x(.),p_2)}{\partial T}=\left[\frac{1}{T}u(x(T))+(p_2-s-g_1)x(T)-c\left(\frac{X}{K}\right)\right]f(T) \tag{56}$$

$$\frac{\partial W(K^e,T^e,x^e(.),p_2^e)}{\partial T}=\frac{F(T^e)}{(T^e)^2}u(x(T^e))>0 \tag{57}$$

We conclude that the equilibrium consumer market size is too small relative to the welfare optimal level.

Similarly, using Eq. (55) to find the partial derivative of welfare with respect to the allocation $x(t)$,

$$\frac{\partial W(K,T,x(.),p_2)}{\partial x(t)}=\int_0^T\left(\left(\frac{1}{t}u'(x(t))+(p_2-s-g_1)\right)f(t)\,\mathrm{d}t\right) \tag{58}$$

$$\frac{\partial W(K^e,T^e,x^e(.),p_2^e)}{\partial x(t)}=\frac{F(t)}{t^2}u'(x^e(t))>0 \tag{59}$$

This shows that for any given type t of consumer, the equilibrium content allocation is too less relative to the social optimum.

Finally, in order to find the partial derivative of welfare with respect to the termination fee p_2, we use Eq. (52)

$$\frac{\partial W(K,T,x(.),p_2)}{\partial p_2} = \frac{\partial Y}{\partial p_2}\left(p_2 - g_1 - \frac{F(T)}{K}c'\left(\frac{Y}{K}\right)\right) - \left(s + Y\frac{\partial s}{\partial Y}\right)\frac{\partial Y}{\partial p_2} + Y \tag{60}$$

Using the first order condition for equilibrium termination fee, that is, Eq. (50), we have

$$\frac{\partial W(K^e,T^e,x^e(.),p_2^e)}{\partial p_2} = 0 \tag{61}$$

That is, the monopolist sets a termination fee that is socially optimal.

We have the following summary result about the welfare properties of equilibrium.

Theorem 4. *With consumers facing nonlinear pricing and content providers facing linear usage pricing, the monopolist service provider's*

(i) *equilibrium provision of capacity is socially optimal.*
(ii) *equilibrium choice of consumer market size in equilibrium is too small relative to the social optimum.*
(iii) *equilibrium usage allocation is too less relative to the social optimum for any given type of the consumer.*
(iv) *equilibrium termination fee p_2^e is socially optimal.*

3.3. Zero-price rule

Finally, we come to the welfare evaluation of a zero-price rule that sets the termination fee $p_2 = 0$. From Eq. (60), we have

$$\begin{aligned}\frac{\partial W}{\partial p_2} &= p_2\frac{\partial Y}{\partial p_2} + Y - \frac{\partial Y}{\partial p_2}\left(e + s + Y\frac{\partial s}{\partial Y}\right)\\ &= p_2\frac{\partial Y}{\partial p_2} + Y\left[1 + \frac{\eta(Y,p_2)}{p_2}\left(e + s\left(1 - \frac{1}{\eta(Y,s)}\right)\right)\right]\end{aligned} \tag{62}$$

The determinants of a marginal change in welfare due to a marginal increase in p_2 are clear from Eq. (62). The first order effect is that the service provider can earn more money. The second order effect is that a higher termination fee reduces content flow on the network and as a result reduces the strength of the first effect; at the same time, it contributes to a reduced marginal social cost of usage and reduced consumer payments towards content usage. There is a third order effect that makes its way through the market for content. A higher termination, by reducing the supply of content flows, increases the market price s for content usage, thereby reducing the strength of the second order effect that reduced consumer payments incurred towards content usage.

Let $(K^0, T^0, x^0(.), Y^0, s^0, e^0)$ denote the equilibrium values of the endogenous variables in the model under a zero-price rule. Then a sufficient condition for $\frac{\partial W(p_2=0)}{\partial p_2} < 0$ is that $e^0 + s^0\left(1 - \frac{1}{\eta^0(Y,s)}\right) < 0$ which is equivalent to the condition that $\eta^0(Y,s) < \frac{s^0}{e^0+s^0}$. Similarly a sufficient condition for $\frac{\partial W(p_2=0)}{\partial p_2} > 0$ is that $1 - \frac{1}{\eta^0(Y,s)} > 0$ which is equivalent to the condition that $\eta^0(Y,s) > 1$. We have the following result.

Theorem 5. *With consumers facing nonlinear pricing and content providers facing linear usage pricing,*

(i) *a sufficient condition for positive termination fee to be socially suboptimal is $\eta^0(Y,s) < \frac{s^0}{e^0+s^0}$.*
(ii) *a sufficient condition for zero termination fee to be socially suboptimal is $\eta^0(Y,s) > 1$.*

Remark. Theorem 5 relates the welfare effects of a zero-price rule to the characteristics of consumer demand for content. If the content flow is sufficiently inelastic with respect to the content price, then a given reduction in content flow leads to a sharp increase in content price. This makes it possible for the third order effect to dominate the first and the second order effects. Imposing a positive termination fee then reduces welfare. On the other hand, if content flow is elastic with respect to the content price, then this ensures that a given change in content flow is met with less than a proportional change in content price making it difficult for the third order effect to dominate. In this case, zero-price rule is suboptimal and a positive termination fee increases welfare.

4. Concluding remarks

The overall modeling framework in this paper is built on some salient economic features of the internet- the complementarity between broadband connection and content, network congestion externalities on the consumer side and oligopolistic externalities on the provider side. We derive equilibrium pricing and investment decisions in the model and study their welfare properties. We display how pricing formulas take into account network congestion and complementarity inherent in the environment. The two models that we study present an interesting contrast about effects of a policy that regulates termination fee. While the policy is neutral in the two-part tariff model, it is not in the model with nonlinear pricing. In the latter model, the privately set termination fee by the monopolist is socially optimal. In spite of this result, we show that the welfare optimality of the neutrality regulation that takes the form of a zero-price rule depends on the elasticity of content demand.

The modeling framework is flexible enough to facilitate the study of many extensions. We point out some here. One direction is to extend the model to a duopoly in internet service provision to study how competitive forces shape the pricing and investment decisions. In this paper, the content market was modeled as a homogenous product symmetric oligopoly. This is an admittedly simplistic view and more realistic formulations can be studied within the modeling framework described here. One can also study the consequences of modeling congestion as affecting the marginal value of content instead of modeling it as a lump-sum tax on utility as we do here.

Declaration of Competing Interest

1. I received no financial support for this research.
2. I received no financial support fromany interested party.
3. I do not hold paid or unpaid positions as officer, director, or board member of relevant non-profit organizations or profit-making entities.
4. The disclosure is true for any close relative or partner of mine.
5. No other party had the right to review the paper prior to its circulation.

Acknowledgments

All the authors declare that they have no relevant material or financial interests related to this research. We thank Felix Munoz–Garcia and the seminar audience at Indian Statistical Institute Delhi for comments.

Appendix 1

The purpose of this appendix is to reformulate the model in Section 2 for the case when the content of different providers is differentiated and the providers strategically set prices. We show that Theorem 1 holds in this reformulated model as well. Aiming for the simplest reformulation, we model the content provider market as a differentiated Bertrand duopoly with linear demand structure. In the interest of brevity, we only spell out the model components that differ from the reference model of Section 2. The ISP chooses two-part tariffs (P_1, p_1) on the consumer side and linear termination fee p_2 on the provider side. There are two CPs, say CP_1 and CP_2, who choose their content prices s_1 and s_2 respectively. Their associated supply of content flows is Y_1 and Y_2, respectively. As a matter of syntax, we assume that the supply of all symbols used in this appendix is available to us afresh even though some of them have been used in Section 2.

Adapting the exposition[22] in Belleflamme and Peitz (2015) of Singh and Vives (1984) to our context, a consumer of type t's utility is

$$U(x_1, x_2, d, P_1, p_1, s_1, s_2; t) = \frac{1}{t}\left[ax_1 + ax_2 - \frac{1}{2}\left(bx_1^2 + 2dx_1x_2 + bx_2^2\right)\right] - c(d) - P_1 - p_1(x_1 + x_2) - s_1x_1 - s_2x_2 \tag{63}$$

if the consumer consumes x_1 and x_2 amounts of content from CP_1 and CP_2, respectively and a, b, d are positive parameters with $b > d$. The latter implies the content from the two providers are imperfect substitutes.

Consumer demand

Consumer of type t faces the following problem of choosing content usage $(x_1(t), x_2(t))$ given the prices

$$\max_{x_1 \geq 0, x_2 \geq 0} \quad \frac{1}{t}\left[ax_1 + ax_2 - \frac{1}{2}\left(bx_1^2 + 2dx_1x_2 + bx_2^2\right)\right] - c(d) - P_1 - p_1(x_1 + x_2) - s_1x_1 - s_2x_2$$
$$\text{subject to} \quad \frac{1}{t}\left[ax_1 + ax_2 - \frac{1}{2}\left(bx_1^2 + 2dx_1x_2 + bx_2^2\right)\right] - c(d) - P_1 - p_1(x_1 + x_2) - s_1x_1 - s_2x_2 \geq 0 \tag{64}$$

The first order conditions for the unconstrained consumer's problem give

$$x_1(t) = \frac{a - tp_1}{b + d} - \frac{tb}{b^2 - d^2}s_1 + \frac{td}{b^2 - d^2}s_2 \tag{65}$$

$$x_2(t) = \frac{a - tp_1}{b + d} + \frac{td}{b^2 - d^2}s_1 - \frac{tb}{b^2 - d^2}s_2 \tag{66}$$

Define

$$\tilde{a} = \frac{a}{b + d}, \quad \tilde{b} = \frac{b}{b^2 - d^2}, \quad \tilde{d} = \frac{d}{b^2 - d^2}, \quad h(T) = \mathbb{E}[t | 0 < t < T]$$

[22] see pp. 65 of the text.

Then $\tilde{a}, \tilde{b}, \tilde{d}$ and $h(T)$ are all positive. The demand for content flows from CP_1 and CP_2 are then given by

$$X_1(s_1, s_2, p_1, T) = \int_0^T x_1(t) f(t)\, \mathrm{d}t = \tilde{a} - h(T)\tilde{b}s_1 + h(T)\tilde{d}s_2 - h(T)(\tilde{b} - \tilde{d})p_1 \tag{67}$$

$$X_2(s_1, s_2, p_1, T) = \int_0^T x_2(t) f(t)\, \mathrm{d}t = \tilde{a} + h(T)\tilde{d}s_1 - h(T)\tilde{b}s_2 - h(T)(\tilde{b} - \tilde{d})p_1 \tag{68}$$

Market clearing in flows

The market clearing condition for content flow now applies not just in the aggregate but individually to content flows from each content provider.

$$X_1 = Y_1 \tag{69}$$

$$X_2 = Y_2 \tag{70}$$

$$X := X_1 + X_2 = Y_1 + Y_2 =: Y \tag{71}$$

Price-setting subgame in the content provider market

The market clearing conditions, Eqs. (69) and (70), imply that demand flows are also supply flows. CP_1's price setting problem is then given by $\max_{s_1} (s_1 - p_2) Y_1(s_1, s_2, p_1, T)$ and analogously for CP_2. The best response functions are

$$s_1 = \frac{\tilde{a} + h(T)\tilde{d}s_2 + h(T)\tilde{b}p_2 - h(T)(\tilde{b} - \tilde{d})p_1}{2h(T)\tilde{b}}, \quad s_2 = \frac{\tilde{a} + h(T)\tilde{d}s_1 + h(T)\tilde{b}p_2 - h(T)(\tilde{b} - \tilde{d})p_1}{2h(T)\tilde{b}}$$

Solving this gives a symmetric price equilibrium of

$$s_1 = s_2 = s = \frac{\tilde{a} + h(T)\tilde{b}p_2 - h(T)(\tilde{b} - \tilde{d})p_1}{h(T)(2\tilde{b} - \tilde{d})} \tag{72}$$

Using individual demand flow Eqs. (67) and (68) and market clearing conditions, Eqs. (69) and (70), we find in (73) that the aggregate supply flow on ISP's communication channels depends only on the total price level $p = p_1 + p_2$ and not the price structure.

$$\begin{aligned} Y = Y_1 + Y_2 &= 2\tilde{a} - 2h(T)(\tilde{b} - \tilde{d})(p_1 + s) \\ &= \frac{2\tilde{a}\tilde{b}}{2\tilde{b} - \tilde{d}} - \frac{2\tilde{b}(\tilde{b} - \tilde{d})h(T)}{2\tilde{b} - \tilde{d}} p \qquad \text{(using (72))} \end{aligned} \tag{73}$$

Service provider's capacity choice and price leadership problem

The service provider's problem in a subgame perfect equilibrium is

$$\max_{K, P_1, p_1, p_2} \quad P_1 F(T) + p_1 X + p_2 Y - g_0(K) - g_1 Y$$

$$\begin{aligned} \text{subject to} \quad & \text{Market Clearing as specified by Equation (71)} \\ & \text{Participation Constraints for all consumers as specified by Equation (64)} \end{aligned}$$

Let $v(p_1, s_1, s_2, T) = \frac{1}{T}\Big[ax_1(T) + ax_2(T) - \frac{1}{2}\Big(b(x_1(T))^2 + 2dx_1(T)x_2(T) + b(x_2(T))^2\Big)\Big] - p_1(x_1(T) + x_2(T)) - s_1x_1(T) - s_2x_2(T)$. Now from the participation constraint (64) for the marginal consumer's problem, it is clear that the service provider will set an access price P_1 so that

$$v(p_1, s_1, s_2, T) = c\left(\frac{X(s_1, s_2, p_1, T)}{K}\right) + P_1 \tag{74}$$

The existence and uniqueness of a marginal type T such that the above equation holds is guaranteed due to the same reasons as in Section 2.

Using Eq. (74) to substitute for optimal access price P_1, the market clearing postulate $X = Y$, and the dependence of supply flows Y only on usage price level p and not on the usage price structure, we may rewrite the ISP's problem in an equilibrium as

$$\max_{K, T, p} \quad \left[v(p, T) - c\left(\frac{Y(p, T)}{K}\right)\right] F(T) + (p - g_1) Y(p, T) - g_0(K)$$

This problem is quite the same as what the monopolist ISP solves in the model of Section 2. Of course, the meaning of symbols is with respect to the reformulated model but conceptually there is little change. Therefore the principal conclusion derived from it - Theorem 1 - holds in the present reformulation. In other words, the conclusions of Theorem 1 does not depend on whether we model the content provider market as a homogenous product Cournot market or as a differentiated product Bertrand market (at least when modeled as a duopoly with linear demand structure). We omit the analysis of nonlinear pricing here in this version of the model.

Appendix 2

Derivation of the inverse demand curve in the non-linear pricing model

The derivation closely follows the textbook exposition in Börgers (2015) and is included here because the presence of other aspects of the model lead to expressions being less than straightforward and omitting the derivation negatively impacts the clarity.

Lemma 1. *If the direct mechanism $(x(.), \mathfrak{p}_1(.))$ is incentive compatible, then the allocation rule $x(.)$ is decreasing in t.*

Proof. Consider two types t and t' with $t > t'$. Incentive compatibility requires that type t does not gain from pretending to be type t' and vice versa. These conditions can be written as

$$\frac{1}{t}u(x(t)) - c\left(\frac{X}{K}\right) - \mathfrak{p}_1(t) - sx(t) \geq \frac{1}{t}u(x(t')) - c\left(\frac{X}{K}\right) - \mathfrak{p}_1(t') - sx(t')$$
$$\frac{1}{t'}u(x(t)) - c\left(\frac{X}{K}\right) - \mathfrak{p}_1(t) - sx(t) \leq \frac{1}{t'}u(x(t')) - c\left(\frac{X}{K}\right) - \mathfrak{p}_1(t') - sx(t')$$

Subtracting the second inequality from the first and invoking the strict monotonicity of $u(.)$ establishes the result. □

Lemma 2. *If the direct mechanism $(x(.), \mathfrak{p}_1(.))$ is incentive compatible, then the payoff function $U(t)$ is decreasing and convex in t. So $U(t)$ is differentiable except at countably many points. For all t at which it is differentiable, $U'(t) = -\frac{1}{t^2}u(x(t))$.*

Proof. Incentive compatibility is equivalent to the following assertion

$$\forall t \in (0, T], \quad U(t) = \max_{t' \in (0,T]} \frac{1}{t}u(x(t')) - c\left(\frac{X}{K}\right) - \mathfrak{p}_1(t') - sx(t')$$

Given any t', the objective function in the maximization problem above is a decreasing and convex function of t. Therefore, $U(t)$ as the maximum of decreasing and convex functions is decreasing and convex as well. The differentiability statement is a standard result in real analysis. The equation for the derivative follows from the Envelope Theorem. □

Lemma 3. *Consider an incentive compatible direct mechanism. Then*

$$\forall t \in (0, T], \quad U(t) = U(T) + \int_t^T \frac{1}{\theta^2}u(x(\theta))\,\mathrm{d}\theta$$

Proof. Since $U(t)$ is convex, it is absolutely continuous and hence the integral of its derivative. □

Lemma 3 shows that the expected utility of any consumer is pinned down by the allocation rule $x(.)$ and the expected utility $U(T)$ of the highest type of the consumer.

Lemma 4. *Consider an incentive compatible direct mechanism. Then*

$$\forall t \in (0, T], \quad \mathfrak{p}_1(t) = \mathfrak{p}_1(T) + \left[\frac{1}{t}u(x(t)) - \frac{1}{T}u(x(T))\right] - s\Big[x(t) - x(T)\Big] - \int_t^T \frac{1}{\theta^2}u(x(\theta))\,\mathrm{d}\theta$$

Proof. Substitute from Eq. (39) the expression for $U(t)$ and $U(T)$ in Lemma 3 and solve for $\mathfrak{p}_1(t)$. □

Lemma 4 shows that the expected payment by any consumer is completely determined by the allocation rule $x(.)$ and the expected payment $\mathfrak{p}_1(T)$ by the highest type of the consumer. Lemmas 1 and 4 give necessary conditions for the direct mechanism $(x(.), \mathfrak{p}_1(.))$ to be incentive compatible. Proposition 1 is a complete characterization whose proof is standard and theorefore omitted.

Proposition 1. *$(x(.), \mathfrak{p}_1(.))$ is incentive compatible if and only if (i) $x(.)$ is decreasing; and (ii) for every $t \in (0, T]$*

$$\mathfrak{p}_1(t) = \mathfrak{p}_1(T) + \left[\frac{1}{t}u(x(t)) - \frac{1}{T}u(x(T))\right] - s[x(t) - x(T)] - \int_t^T \frac{1}{\theta^2}u(x(\theta))\,\mathrm{d}\theta$$

Proposition 2. *An incentive compatible mechanism $(x(.), \mathfrak{p}_1(.))$ is individually rational if and only if $U(T) \geq 0$.*

Proof. Necessity is obvious. For sufficiency, note by Lemma 2 that $U(t)$ is decreasing in t for incentive compatible mechanisms. So for every $t \in (0, T]$, $U(t) \geq U(T) \geq 0$. □

Lemma 5. *If an incentive compatible and individually rational direct mechanism $(x(.), \mathfrak{p}_1(.))$ maximizes the service provider's expected profits, then*[23]

$$\mathfrak{p}_1(T) = \frac{1}{T}u(x(T)) - c\left(\frac{X}{K}\right) - sx(T)$$

[23] Lemma 5, when viewed as $\frac{1}{T}u(x(T)) - sx(T) = c\left(\frac{X}{K}\right) + \mathfrak{p}_1(T)$, is the analogue of Eq. (21) in this model.

Proof. By Proposition 2, $\mathfrak{p}_1(T) \le \frac{1}{T}u(x(T)) - c\left(\frac{X}{K}\right) - sx(T)$. The service provider's expected profits would be maximum if $\mathfrak{p}_1(T)$ is set at its upper bound. $\square$

Substituting the highest type's payment from Lemma 5 into the payment formula of Proposition 1, we get the important result that the pricing rule is completely determined (up to congestion cost) by the allocation rule

$$\forall t \in (0, T], \quad \mathfrak{p}_1(t) = \frac{1}{t}u(x(t)) - sx(t) - c\left(\frac{X}{K}\right) - \int_t^T \frac{1}{\theta^2}u(x(\theta))\,\mathrm{d}\theta \tag{75}$$

Lemma 5 then gives the inverse demand curve as Eq. (42) in Section 3.

Argument for reduction of [OPT1] to [OPT2]

Using Eq. (75), the first term in the objective function of [OPT1] can be rewritten as

$$\int_0^T \left[\frac{1}{t}u(x(t)) - sx(t) - c\left(\frac{X}{K}\right)\right] f(t)\,\mathrm{d}t - \int_0^T \int_t^T \frac{1}{\theta^2}u(x(\theta))\,\mathrm{d}\theta\, f(t)\,\mathrm{d}t$$

Now consider the second integral in the expression above. By Fubini's Theorem, we can change the order of integration and leave the integral unchanged. That is

$$\begin{aligned}\int_0^T \int_t^T \frac{1}{\theta^2}u(x(\theta)) f(t)\,\mathrm{d}\theta\,\mathrm{d}t &= \int_0^T \int_0^\theta \frac{1}{\theta^2}u(x(\theta)) f(t)\,\mathrm{d}t\,\mathrm{d}\theta \\ &= \int_0^T \frac{1}{\theta^2}u(x(\theta))\left(\int_0^\theta f(t)\,\mathrm{d}t\right)\mathrm{d}\theta \\ &= \int_0^T \frac{1}{\theta^2}u(x(\theta))F(\theta)\,\mathrm{d}\theta = \int_0^T \frac{1}{t^2}u(x(t))F(t)\,\mathrm{d}t\end{aligned}$$

This shows [OPT1] can be reduced to [OPT2].

References

Armstrong, M., 2006. Competition in two-sided markets. RAND J. Econ. 37 (3), 668–691.
Belleflamme, P., Peitz, M., 2015. Industrial Organization: Markets and Strategies. Cambridge University Press.
Börgers, T., 2015. An Introduction to the Theory of Mechanism Design. Oxford University Press, USA.
Cournot, A.-A., 1838. Recherches sur les principes mathématiques de la théorie des richesses par augustin cournot. Chez L. Hachette.
Economides, N., 2007. The economics of the internet. The New Palgrave Dictionary of Economics. Macmillan, London.
Economides, N., Hermalin, B.E., 2012. The economics of network neutrality. RAND J. Econ. 43 (4), 602–629.
Economides, N., Tåg, J., 2012. Network neutrality on the internet: a two-sided market analysis. Inf. Econ. Policy 24 (2), 91–104.
Farrell, J., Saloner, G., 1985. Standardization, compatibility, and innovation. RAND J. Econ. 70–83.
Farrell, J., Saloner, G., 1986. Installed base and compatibility: innovation, product preannouncements, and predation. Am. Econ. Rev. 940–955.
Greenstein, S., 2020. The basic economics of internet infrastructure. J. Econ. Perspect. 34 (2), 192–214.
Greenstein, S., Peitz, M., Valletti, T., 2016. Net neutrality: a fast lane to understanding the trade-offs. J. Econ. Perspect. 30 (2), 127–150.
Hemphill, C.S., 2008. Network neutrality and the false promise of zero-price regulation. Yale J. Reg. 25, 135.
Katz, M.L., Shapiro, C., 1985. Network externalities, competition, and compatibility. Am. Econ. Rev. 75 (3), 424–440.
Katz, M.L., Shapiro, C., 1986. Technology adoption in the presence of network externalities. J. Polit. Econ. 94 (4), 822–841.
Lee, R.S., Wu, T., 2009. Subsidizing creativity through network design: zero-pricing and net neutrality. J. Econ. Perspect. 23 (3), 61–76.
MacKie-Mason, J.K., Varian, H.R., 1995. Pricing congestible network resources. IEEE J. Sel. Areas Commun. 13 (7), 1141–1149.
Maskin, E., Riley, J., 1984. Monopoly with incomplete information. RAND J. Econ. 15 (2), 171–196.
McManus, M., 1962. Numbers and size in Cournot oligopoly. Bull. Econ. Res. 14 (1), 14–22.
McManus, M., 1964. Equilibrium, numbers and size in Cournot oligopoly. Bull. Econ. Res. 16 (2), 68–75.
Mussa, M., Rosen, S., 1978. Monopoly and product quality. J. Econ. Theory 18 (2), 301–317.
Oi, W.Y., 1971. A disneyland dilemma: two-part tariffs for a mickey mouse monopoly. Q. J. Econ. 85 (1), 77–96.
Pil Choi, J., Kim, B.-C., 2010. Net neutrality and investment incentives. RAND J. Econ. 41 (3), 446–471.
Roberts, J., Sonnenschein, H., 1976. On the existence of Cournot equilbrium without concave profit functions. J. Econ. Theory 13 (1), 112–117.
Rochet, J.-C., Tirole, J., 2003. Platform competition in two-sided markets. J. Eur. Econ. Assoc. 1 (4), 990–1029.
Rochet, J.-C., Tirole, J., 2006. Two-sided markets: a progress report. RAND J. Econ. 37 (3), 645–667.
Schmalensee, R., 1981. Monopolistic two-part pricing arrangements. Bell J. Econ. 445–466.
Shapiro, C., Varian, H.R., 1998. Information Rules: A Strategic Guide to the Network Economy. Harvard Business Press.
Singh, N., Vives, X., 1984. Price and quantity competition in a differentiated duopoly. RAND J. Econ. 546–554.
Varian, H.R., 1989. Price discrimination. In: Handbook of Industrial Organization, 1. Elsevier, pp. 597–654.
Wu, T., 2003. Network neutrality, broadband discrimination. J. Telecommun. High Technol. Law 2, 141.