An Economic Model of the Last-Mile Internet

Rakesh Chaturvedi *
Souvik Dutta †
Kiriti Kanjilal ‡

Abstract

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

JEL Classification: D42, D43, D46, L11, L12, L13, L86, L96
Keywords: internet, network neutrality, congestion pricing, two-part tariffs, nonlinear pricing

1 Introduction

The commercial internet is a ubiquitous presence in people's lives today and supports a wide range of activities in various domains like business, media, leisure and education. Today, the majority of data traffic on the internet comes from video-streaming and gaming applications unlike the early days when it primarily came from email and file transfer applications. Structurally, the global internet is a network of networks that is supported by physical infrastructure like root servers, fiber/broadband lines, network switches and routers, content delivery networks, cellular towers etc. in addition to the interconnecting backbone lines. Greenstein (2020) and Economides (2007) are recent surveys that describes many aspects of the internet infrastructure and some related economic questions.

*Department of Social Sciences and Humanities, Indraprastha Institute of Information Technology Delhi
†Department of Social Sciences and Humanities, Indraprastha Institute of Information Technology Delhi
‡Department of Social Sciences and Humanities, Indraprastha Institute of Information Technology Delhi

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 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 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[1] 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* [2] fees to non-originating content providers [3] 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.

A formal evaluation of pricing and investment policies concerning the internet needs a model of the internet ecosystem. Our paper is a contribution in that direction. We take some economic aspects of the internet ecosystem as fundamental. The first of these is the complementarity of broadband connection and content for the consumers. 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. Another economic aspect is the separation of ownership between providers of the internet connection and providers of content. In this paper, we model the market for last-mile internet service provision as a monopoly which makes investment and pricing decisions on two sides- the consumer side and the content provider side. We then impose a neutrality regulation that takes the form of a zero-price rule to evaluate its welfare implications. On the provider side, we posit a simple oligopolistic market structure for supplying content resulting in production externalities. The market in content is modeled as an entry-augmented Cournot game of quantity competition in substitute goods. 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 a coherent integration of these two models that casts the internet in a demand-supply framework. On the consumer side, there are congestion externalities resulting from

[1] 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.

[2] 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.

[3] 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.

aggregate demand. We borrow the quasilinear consumer utility specification (including a specification of congestion costs) from MacKie-Mason and Varian (1995) which provides a simple microfoundation for consumer demand in our model. Our model, compactly depicted in Figure 1, presents a simple framework to study pricing of data flows on the internet and investment incentives for the service provider.

We first study the baseline model in which consumers 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 its composition on the consumer and the content provider side. The usage price level is given by a congestion-adjusted Lerner-style formula. 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 expository model. The neutrality result prompts us to extend our baseline model to one in which consumers who are heterogenous in tastes face nonlinear prices from the service provider. We derive equilibrium prices on both sides of the market. While the equilibrium termination fee is given by a congestion-adjusted Lerner-style formula, 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 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 monopoly. In the extended model with nonlinear pricing, equilibrium price on the consumer side is too high relative to the social optimum. On the provider side, the equilibrium termination fee is too high with respect to the socially optimal level if and only if the elasticity of consumer demand with respect to content usage rice is sufficiently low. 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 and only if the consumer demand for content is sufficiently convex. 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), Rochet and Tirole (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, 1986) and Katz and Shapiro (1985, 1986)) or cross group positive externalities that are characteristic of these markets are indirect and work through endogenous channels. In our model, a consumer wants more providers because competition can drive down the price of content. Similarly, a provider cares about more consumers because he can look forward to more sales. 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 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 pure welfare analysis and equilibrium analysis follow in Sections 2.1 and 2.2. Welfare properties of 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 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 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 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

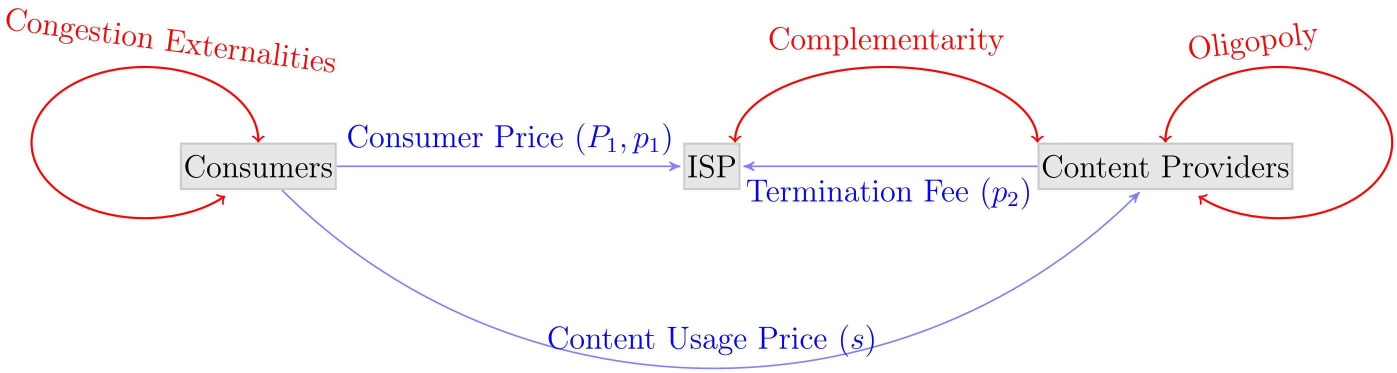

Figure 1: Model Structure. The acronym ISP stands for internet service provider.The entities in the model are in black. The blue connections indicate the flow of money. The red connections point out the salient economic features.

channels. It is understood that any given content possibly traverses many networks before reaching the consumer through the last-mile provider.

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, therefore, is a model in which content from various providers are substitutes. 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 to the utilization rate. We follow MacKie-Mason and Varian (1995) by defining the utilization rate as a simple ratio of total usage over capacity X/K and setting the congestion delay d equal to the utilization rate, that is, $d = X/K$. Figure 1 shows the salient features of the model in a compact representation.

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.

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

where (P_1, p_1) is the two-part tariff 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 F 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 [4]. 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 $j = 1, \dots, J$, we write content provider j's profit as

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

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. Moreover, we 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. Figure 2 provides an illustration of content flows in the model. The market clearing (in flows) condition is given by Equation (3) where T is the marginal consumer who connects to the service provider.

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

We augment this Cournot supply stage with an entry stage at which potentially large number of providers make their entry decisions.

Let $f(K, X)$ denote the service provider's cost of supporting an aggregate usage X through an installed capacity K. We assume that $f(K, X) = f_0(K) + f_1 X$ where $f_0(K)$ is the fixed cost of installed capacity K and $f_1 > 0$ is the constant marginal cost of aggregate usage. The service provider's profit when T is the marginal consumer is given by

$$\begin{aligned} \pi_I(P_1, p_1, p_2, (y_j)_{j \in J}) &= P_1 F(T) + p_1 X + p_2 \sum_{j \in J} y_j - f(K, X) \\ &= P_1 F(T) + p_1 X + p_2 Y - f_0(K) - f_1 X \end{aligned} \tag{4}$$

[4]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.

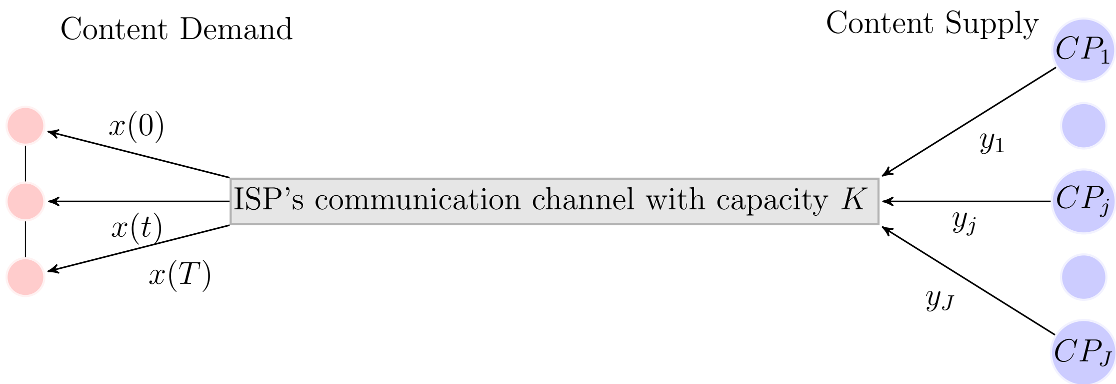

Figure 2: An illustration of content flows in the model.

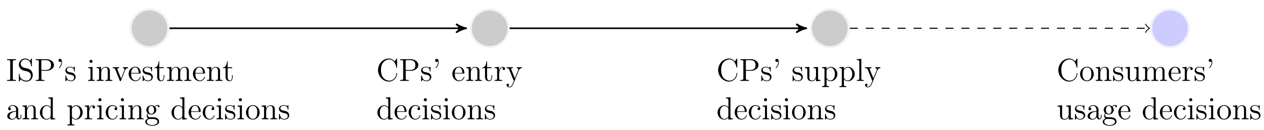

Figure 3: Timeline of decisions in the model. The acronyms ISP and CP stand for internet service provider and content provider respectively. Consumers are modeled as nonstrategic and so theirs is not a decision node in a strict sense. A different node color and a dashed edge serve as reminders of this.

We posit a multi-stage game formulation which is schematically shown in Figure 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 Equation (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 Equation (2). We will study sequential 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

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

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

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

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 $x(t) = (u')^{-1}\big(t(p_1 + s)\big)$. 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

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

The consumer demand exhibits a complementarity relationship between broadband connection and content. From Equation (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

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

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

2.1 Welfare Optimum

In this subsection, we carry out a pure 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. The analysis has aspects of MacKie-Mason and Varian (1995) in spite of our model being more involved. The differences arise from the presence of a continuum and heterogeneity of consumers and marginal service costs.

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

$$\max_{K,T,x(t)} \quad W(K,T,x(t)) = \int_0^T \frac{1}{s} u(x(s)) f(s)\,\mathrm{d}s - F(T) c\Big(\frac{X(T)}{K}\Big) - f_1 X(T) - f_0(K)$$
$$\text{subject to } X(T) = \int_0^T x(s) f(s)\,\mathrm{d}s$$
$$\forall s \in (0,T], \quad x(s) \geq 0$$

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 $(K, T, x(t))$ that solves the following first order conditions of the welfare maximization problem that may be written as

$$\frac{\partial W}{\partial K} = F(T)\frac{1}{K} c'\Big(\frac{X(T)}{K}\Big)\frac{X(T)}{K} - f_0'(K) = 0 \tag{9}$$
$$\frac{\partial W}{\partial T} = \frac{1}{T} u(x(T)) f(T) - f(T) c\Big(\frac{X(T)}{K}\Big) - F(T)\frac{1}{K} c'\Big(\frac{X(T)}{K}\Big) x(T) f(T) - f_1 x(T) f(T) = 0 \tag{10}$$
$$\frac{\partial W}{\partial x(t)} = \frac{1}{t} u'(x(t)) f(t) - F(T)\frac{1}{K} c'\Big(\frac{X(T)}{K}\Big) f(t) - f_1 f(t) = 0 \tag{11}$$

From Equation (10), the social value of increasing T is the marginal consumer's value of content. The social cost is a 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

$$\frac{1}{T} u(x(T)) = c\Big(\frac{X(T)}{K}\Big) + F(T)\frac{1}{K} c'\Big(\frac{X(T)}{K}\Big) x(T) + f_1 x(T) \tag{12}$$

On similar lines, from Equation (11), 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

$$\frac{1}{t} u'(x(t)) = F(T)\frac{1}{K} c'\Big(\frac{X(T)}{K}\Big) + f_1 \tag{13}$$

The first term on the right hand side in Equation (13) 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

$$e = p_c + f_1 \tag{14}$$

where $p_c = \frac{\partial c}{\partial X} F(T) = F(T) \frac{1}{K} c' \left(\frac{X(T)}{K} \right)$ 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.

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 Equation (9) in a more interpretable form as

$$\frac{\partial W}{\partial K} = p_c \frac{X(T)}{K} - f_0'(K) = 0 \tag{15}$$

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 that prices adjust to clear the market. Using Equation (3) in Equation (8) gives the residual inverse demand curve for content as

$$s(Y, T, p_1) := \frac{1}{T} u' \Big(\frac{1}{f(T)} \frac{\partial Y}{\partial T} \Big) - p_1 =: \phi(Y, T) - p_1 \tag{16}$$

Equation (16) 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 $p = p_1 + p_2$. 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.

$$s - p_2 = \phi(Y, T) - p \tag{17}$$

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 Equation (2) can be rewritten as $\pi_C^j(P_1, p_1, p_2, y_j, Y) = (s(Y, T, p_1) - p_2) y_j - F$. Ignoring the participation constraints, the corresponding first order conditions are

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

$$\text{Summing over } j, \quad Y \frac{\partial s}{\partial Y} = J(p_2 - s(Y, T, p_1)) \tag{18}$$

Using Equation (17), we can rewrite Equation (18) as

$$Y \frac{\partial \phi}{\partial Y} = J(p - \phi(Y, T)) \tag{19}$$

So the aggregate content supply $Y(p, T, J)$ is determined by Equation (19) if $p_2 \leq s$ and is equal to 0 if $p_2 > s$. 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 run revenue losses. It is clear that the equilibrium in this Cournot subgame [5] is symmetric- 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 profits of a content provider.

Entry Subgame

When there are no barriers to entry other than the fixed costs F, 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).

$$\pi_C(J) = (s - p_2)\frac{Y}{J} - F = 0 \tag{20}$$

Equations (17) and (20) 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

$$J = \frac{(\phi(Y) - p)Y}{F} \tag{21}$$

where $Y(p, T, J)$ is written as Y with its arguments suppressed and is determined in the Cournot subgame by Equation (19). Note that 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 Equation (21) in Equation (19) and is determined from

$$\frac{\partial \phi(Y)}{\partial Y} = -\frac{(\phi(Y) - p)^2}{F} \tag{22}$$

Equation (22) makes it clear that the aggregate content flow is insensitive to the way the usage price level 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. Moreover, it must ensure a positive margin for the content providers to ensure their operational sustainability. This is the participation constraint for the content providers. In

[5] 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).

the present context, we refer to it as the content providers' viability constraint. The service provider's problem in a sequential equilibrium is

$$\begin{aligned}\max_{K,P_1,p_1,p_2} \quad & P_1F(T)+p_1X(p_1,s,T)+p_2Y(p_2,s)-f_0(K)-f_1Y\\ \text{subject to} \quad & \text{Market Clearing as specified by Equation (3)}\\ & \text{Participation Constraints for all consumers as specified by Equation (5)}\\ & \text{Aggregate Demand as specified by Equation (8)}\\ & \text{Aggregate Supply as specified by Equation (22); and}\\ & \text{Content Providers' Viability Constraint as specified by } p_2 \le s\end{aligned}$$

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. 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.

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

$$v(p_1+s,T) = c\Big(\frac{X(p_1+s,T)}{K}\Big) + P_1 \tag{23}$$

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 $c\Big(\frac{X(p_1+s,T)}{K}\Big)$ is an increasing function of T with $\lim_{T\to 0} c\Big(\frac{X(p_1+s,T)}{K}\Big) = 0$. All consumers of type $t \le T$ connect because for any such type $v(p_1+s,t) \ge v(p_1+s,T)$.

Using equation (23) to substitute for optimal access price P_1 and the market clearing postulate $X = Y$, we may rewrite the service provider's problem as that of choosing the capacity K, the marginal consumer T and the total usage price p

$$\begin{aligned}\max_{K,T,p} \quad & \Big[v(p,T) - c\Big(\frac{Y(p,T)}{K}\Big)\Big]F(T) + (p-f_1)Y(p,T) - f_0(K)\\ \text{subject to} \quad & p \le \phi(Y(p,T),T)\end{aligned}$$

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 Schmalense (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) that strictly satisfies content providers' viability constraint must satisfy the following necessary first order conditions for optimality.

$$\frac{\partial \pi_I}{\partial K} = \frac{F(T)}{K} c'\Big(\frac{Y}{K}\Big)\frac{Y}{K} - f_0'(K) = 0 \tag{24}$$

$$\frac{\partial \pi_I}{\partial T} = \Big[\frac{\partial v(p,T)}{\partial T} - \frac{\partial c}{\partial T}\Big] F(T) + [v(p,T) - c(p,T)]f(T) + (p - f_1)\frac{\partial Y}{\partial T} = 0 \tag{25}$$

$$\frac{\partial \pi_I}{\partial p} = \Big[\frac{\partial v(p,T)}{\partial p} - \frac{\partial c}{\partial p}\Big] F(T) + (p - f_1)\frac{\partial Y}{\partial p} + Y(p,T) = 0 \tag{26}$$

The optimality condition (24) with respect to K is exactly the same as the optimality condition (9) 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. 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' viability constraint.

The optimality conditions with respect to T can be intuitively understood by rewriting it in terms of the access price using Equation (23) which says $v - c = P_1$. The optimality condition now reads

$$\frac{\partial P_1}{\partial T} F(T) + P_1 f(T) + (p - f_1)\frac{\partial Y}{\partial T} = 0 \tag{27}$$

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 consumers' value content less than existing consumers. Moreover, they increase congestion on the network. This reduces the access price that the monopolist can charge.

Using Roy's Identity[6], we have $\frac{\partial v}{\partial p} = -x(p,T)$; 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 Equation (14), the condition with respect to p looks like

$$\frac{p - e}{p} = \frac{1}{\eta(Y,p)}\Big(1 - \frac{x(p,T)}{X(p,T)/F(T)}\Big) \tag{28}$$

Equation (28) resembles the first-order order condition for optimal usage price as developed in Varian (1989). The difference is that here, the monopolist takes into account the congestion costs. 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 e for a social planner which in turn is precisely the marginal social cost of usage $p_c + f_1$ from Equation (14). We close this section with a summary result

[6] 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 $x(p,t) = -\frac{\partial v(p,t)}{\partial p}$.

about the salient features of equilibrium in the model.

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

(i) the monopolist service provider's equilibrium 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 equilibrium usage price level p must satisfy Equation (28), 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 Equation (25) 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:

$$W(K,T,p) = \int_0^T v(p,t)f(t)\,\mathrm{d}t - c\Big(\frac{Y}{K}\Big)F(T) + (p-f_1)Y(p,T) - f_0(K) + (s-p_2)Y(p,T) - JF \tag{29}$$

At the free-entry symmetric sequential equilibrium, $Y = Jy(J)$ and $\pi_C(J) = 0$. So Equation (29) may be rewritten as

$$\begin{aligned} W(K,T,p) &= \int_0^T v(p,t)f(t)\,\mathrm{d}t - c\Big(\frac{Y}{K}\Big)F(T) + (p-f_1)Y(p,T) - f_0(K) + J\pi_C(J) \\ &= \int_0^T v(p,t)f(t)\,\mathrm{d}t - c\Big(\frac{Y}{K}\Big)F(T) + (p-f_1)Y(p,T) - f_0(K) \end{aligned} \tag{30}$$

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 usage price 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

$$\frac{\partial W(K,T,p)}{\partial K} = \frac{F(T)}{K} c'\Big(\frac{Y}{K}\Big)\frac{Y}{K} - f_0'(K) \tag{31}$$
$$\frac{\partial W(K,T,p)}{\partial T} = \Big[v(p,T) - c\Big(\frac{Y}{K}\Big) + (p - f_1)x(p,T)\Big] f(T) \tag{32}$$
$$\begin{aligned}\frac{\partial W(K,T,p)}{\partial p} &= -\int_0^T x(p,t)f(t)\,\mathrm{d}t - \frac{F(T)}{K}c'\Big(\frac{Y}{K}\Big)\frac{\partial Y}{\partial p} + (p - f_1)\frac{\partial Y}{\partial p} + Y(p,T) \\ &= \frac{\partial Y}{\partial p}\Big(p - f_1 - \frac{F(T)}{K}c'\Big(\frac{Y}{K}\Big)\Big)\end{aligned} \tag{33}$$

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 Equations (24), (25) and (26), we have

$$\frac{\partial W(K^e,T^e,p^e)}{\partial K} = 0 \tag{34}$$
$$\frac{\partial W(K^e,T^e,p^e)}{\partial T} = \Big[\frac{\partial c(p^e,T^e)}{\partial T} - \frac{\partial v(p^e,T^e)}{\partial T}\Big] F(T^e) > 0 \tag{35}$$
$$\frac{\partial W(K^e,T^e,p^e)}{\partial p} = x(p^e,T^e)F(T^e) - X(p^e,T^e) \tag{36}$$

From Equation (34), we conclude that the equilibrium provision of capacity by the service provider is socially optimal. From Equation (35), 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 Equation (36), we have

$$\frac{\partial W(K^e,T^e,p^e)}{\partial p} > 0 \quad \text{if and only if} \quad x(p^e,T^e) - \frac{X(p^e,T^e)}{F(T^e)} > 0 \tag{37}$$

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 $\pi_I(p_1, p_2, (y_j)_{j\in J}) = p_1 X + p_2 \sum_{j\in J} y_j - f(K, X)$. 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

$$u(X) - c\Big(\frac{X}{K}\Big) - (p_1 + s)X \geq 0 \tag{38}$$

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

$$p_1 + s = u'(X) \tag{39}$$

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

$$u''(Y) = -\frac{(u'(Y) - p)^2}{F} \tag{40}$$

when $p_2 \leq s$ and equals 0 otherwise. The aggregate content supply is insensitive to the service provider's price structure (p_1, p_2) and is sensitive only to the price level $p = p_1 + p_2$. This means that in the special case of this example, the resulting market is not a two-sided market in the sense of Rochet and Tirole (2006). The service provider's problem in a sequential equilibrium is

$$\begin{aligned} \max_{K, p_1, p_2} \quad & p_1 X(s, p_1) + p_2 Y(s, p_2) - f_0(K) - f_1 Y \\ \text{subject to} \quad & \text{Market Clearing as specified by Equation (3)} \\ & \text{Aggregate Demand as specified by Equation (39)} \\ & \text{Aggregate Supply as specified by Equation (40)} \\ & \text{Universal Service as specified by Equation (38); and} \\ & \text{Content Providers' Viability Constraint as specified by } p_2 \leq s \end{aligned}$$

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

$$\begin{aligned} \max_{K, p} \quad & (p - f_1)Y(p) - f_0(K) \\ \text{subject to} \quad & p \leq u'(Y) \ \text{ and Equation (38)} \end{aligned}$$

The solution to this problem is economically interesting only when both the constraints is slack at the optimum. Ignoring the constraints, this is the standard monopoly problem and the optimal price level is determined by the standard markup formula

$$\frac{p - f_1}{p} = \frac{1}{\eta} \tag{41}$$

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 Equation (38).

3 Consumers facing nonlinear pricing from service provider

In this section, we generalize the 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

$$U(t) = \frac{1}{t}u(x(t)) - c(d) - \mathfrak{p}_1(t) - sx(t) \tag{42}$$

where symbols inherit the meaning they held in Section 2. The service provider's profit when T is the marginal consumer is given by

$$\pi_I(\mathfrak{p}_1(.), p_2, (y_j)_{j \in J}) = \int_0^T \mathfrak{p}_1(t) f(t)\, \mathrm{d}t + p_2 \sum_{j \in J} y_j - f(K, X) \tag{43}$$

The market clearing condition now looks like

$$X(\mathfrak{p}_1(.), s, T) = Y(s, p_2) \tag{44}$$

Given the type-contingent content allocation $x(.)$, the aggregate demand for content when T is the marginal consumer is given by

$$X(\mathfrak{p}_1(.), s, T) = \mathbb{E}[x(t) | t \in (0, T)] = \int_0^T x(t) f(t)\, \mathrm{d}t \tag{45}$$

From Equation (45) and the First Fundamental Theorem of Calculus, we get $x(T)f(T) = \frac{\partial X}{\partial T}$. The welfare analysis of this model is the same as in Section 2.1 and is omitted here.

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 models of screening. Nevertheless, we derive it in the context of the present model in which nonlinear pricing is only one aspect. The derivation closely follows the textbook exposition in Börgers (2015) (see Mussa and Rosen (1978), Maskin and Riley (1984) and references therein for a sample of the extensive literature on nonlinear pricing) 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

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

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

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

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

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. Q.E.D.

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. Q.E.D.

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*

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

Proof. Substitute from Equation 42 the expression for $U(t)$ and $U(T)$ in Lemma 3 and solve for $\mathfrak{p}_1(t)$. Q.E.D.

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. Lemma 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* $t \in (0,T]$

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

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 $t \in (0,T]$, $U(t) \geq U(T) \geq 0$. Q.E.D.

Lemma 5. *If an incentive compatible and individually rational direct mechanism* $(x(.), \mathfrak{p}_1(.))$ *maximizes the service provider's expected profits, then*[7]

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

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

Substituting the highest type's payment from Lemma 5 into the payment formula of Proposition 1, we have

$$\forall t \in (0,T], \quad \mathfrak{p}_1(t) = \frac{1}{t}u(x(t)) - sx(t) - c\Big(\frac{X}{K}\Big) - \int_t^T \frac{1}{\theta^2}u(x(\theta))\,\mathrm{d}\theta \tag{46}$$

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

Equation (46) gives the pricing rule (up to the congestion cost) in terms of the allocation rule. Lemma 5 gives the inverse demand curve as

$$s = \frac{1}{T}u'(x(T)) - \frac{\partial \mathfrak{p}_1(T)}{\partial x(T)} \tag{47}$$

This derivation of inverse demand curve is consistent with what was obtained earlier when the model was specialized to the case of consumers facing two-part tariffs. In that case, $\mathfrak{p}_1(T) = P_1 + p_1 x(T)$; so Equation (47) reduces to Equation (8). As in the model with two-part tariffs, 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

Cournot Subgame for Content Supply

Equation (47), the market clearing condition and and an application of the First Fundamental Theorem of Calculus to Equation (45) imply the residual inverse demand curve for content is

$$s(Y, T, \mathfrak{p}_1(T)) = \frac{1}{T}u'\Big(\frac{1}{f(T)}\frac{\partial Y}{\partial T}\Big) - \frac{\partial \mathfrak{p}_1(x(T) = \frac{1}{f(T)}\frac{\partial Y}{\partial T})}{\partial x(T)} \tag{48}$$

All content providers are alike. Following the analysis of Section 2.2, we can find the content supply $Y(\mathfrak{p}_1(T), p_2, T, J)$ for any given number J of content providers from

$$Y\frac{\partial s}{\partial Y} = J\big(p_2 - s(Y, T, \mathfrak{p}_1(T))\big) \tag{49}$$

Entry Subgame

The zero profit condition in this model is

$$\pi_C(J) = (s - p_2)\frac{Y}{J} - F = 0 \tag{50}$$

Equation (50) implies that given the service provider's strategy choice of $(K, T, \mathfrak{p}_1(.), p_2)$, the equilibrium entry J is determined from

$$J = \frac{(s - p_2)Y}{F} \tag{51}$$

The aggregate supply $Y(\mathfrak{p}_1(T), p_2, T)$ of content in equilibrium is determined by substituting Equation (51) in Equation (49) and is determined from

$$\frac{\partial s(Y)}{\partial Y} = -\frac{(s(Y) - p_2)^2}{F} \tag{52}$$

Service Provider's Capacity Choice and Price Leadership Problem

The service provider's problem in a sequential equilibrium is

$$\begin{aligned}
\max_{K,T,\mathfrak{p}_1(.),p_2} \quad & \int_0^T \mathfrak{p}_1(t) f(t)\,\mathrm{d}t + (p_2 - f_1) Y(\mathfrak{p}_1(T), p_2, T) - f_0(K) \\
\text{subject to} \quad & (x(.), \mathfrak{p}_1(.)) \text{ is incentive compatible and individually rational} \\
& \text{Market Clearing as specified by Equation}(44) \\
& \text{Aggregate Demand as specified by Equation}(47) \\
& \text{Aggregate Supply as specified by Equation}(52);\ \text{and} \\
& \text{Content Providers' Viability Constraint as specified by } p_2 \leq s
\end{aligned}$$

Using Equation 46, the first term in the objective function can be rewritten as

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

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

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

We may now frame the service provider's problem as that of choosing the allocation rule $x(.)$ instead of the pricing rule $\mathfrak{p}_1(.)$.

$$\begin{aligned}
\max_{K,T,x(.),p_2} \quad & \int_0^T \Big[\Big(\frac{1}{t} - \frac{F(t)}{t^2}\Big) u(x(t)) - s x(t) - c\Big(\frac{X}{K}\Big)\Big] f(t)\,\mathrm{d}t + (p_2 - f_1) Y(\mathfrak{p}_1(T), p_2, T) - f_0(K) \\
\text{subject to} \quad & x(.) \text{ is decreasing} \\
& \text{Equations } (44), (47), (52);\ \text{and } p_2 \leq s
\end{aligned} \tag{53}$$

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

$$\begin{aligned}
\max_{K,T,x(.)} \quad & \int_0^T \Big[\Big(\frac{1}{t} - \frac{F(t)}{t^2}\Big) u(x(t)) - s x(t) - c\Big(\frac{X}{K}\Big)\Big] f(t)\,\mathrm{d}t + (p_2 - f_1) \int_0^T x(t) f(t)\,\mathrm{d}t - f_0(K) \\
\text{subject to} \quad & x(.) \text{ is decreasing}
\end{aligned} \tag{54}$$

Optimal Allocation Rule

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

$$\Big(\frac{1}{t} - \frac{F(t)}{t^2}\Big)u'(x(t)) = f_1 + s - p_2 \tag{55}$$

Under the constraint that the markup of content providers must be positive, the right side of Equation (55) 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* $T \leq \bar{T}$.

The left side of Equation 55 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 0 for a consumer whose type t approaches 0 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 Equation (55) is positive and tends to infinity when $x(t)$ tends to 0; 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 [8].

Equation 55 can be conveniently restated in words as follows

$$\text{Virtual Marginal Benefit of Type-}t\text{ Consumer} - \text{Content Providers' Markup} = \text{Service Provider's Marginal Costs} \tag{56}$$

The optimal allocation rule in the present model as stated in words by Equation (56) 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

Using Leibniz Rule, we get the optimality condition for T as follows

$$-\frac{F(T)}{T^2}u(x(T)) + \Big(\frac{1}{T}u(x(T)) - sx(T) - c\Big(\frac{X}{K}\Big)\Big)f(T) + (p_2 - f_1)x(T)f(T) = 0 \tag{57}$$

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

The intuitive meaning of this equation can be uncovered by making use of Equation (46) to rewrite Equation (57) as

$$\frac{\partial \mathfrak{p}_1(t=T)}{\partial T} F(T) + \mathfrak{p}_1(T) f(T) + (p_2 - f_1) \frac{\partial X}{\partial T} = 0 \tag{58}$$

Equation (58) establishes consistency of this optimality condition with its counterpart, Equation (27), in the two-part tariff model. At the same time, it reveals that the tradeoffs involved here are no different than that model. The optimal market size is one that optimally resolves the tradeoff of a marginal increase in revenues of the service provider on account of its ability to price the increased content flows on both sides with the marginal decrease in consumer prices that must follow if the market size increases.

Optimal Capacity Investment

Using the definition of p_c in Equation (14) as the marginal social cost of congestion, the optimal capacity investment solves the first order condition

$$p_c X = f_0'(K) K \tag{59}$$

Optimal Termination Fee

When looking to derive optimality conditions for p_2, we use the formulation in 53. With the definition of Pigouvian tax e from Equation (14), the optimality condition resembles a Lerner's formula. As in the two-part tariff model, the markup is over the Pigouvian tax (marginal social cost of usage) that is set so consumers internalize the congestion externalities.

$$\frac{p_2 - e}{p_2} = \frac{1}{\eta\big(Y(\mathfrak{p}_1(T), p_2, T), p_2\big)} \tag{60}$$

where $\eta\big(Y(\mathfrak{p}_1(T), p_2, T), p_2\big)$ is the elasticity of content supply with respect to 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. 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.

We have the following summary result about salient features of 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 congestion cost) by the equilibrium usage allocation rule $x(.)$ for consumers through Equation (46). Under Assumption 1 and 2, the equilibrium usage allocation rule $x(.)$ must satisfy Equation (55) [9]*.*

[9] see Equation (56) for an intuitive interpretation.

(ii) equilibrium termination fee p_2 must satisfy Equation (60), a Lerner-style formula in which the markup is over the Pigouvian tax.

(iii) equilibrium consumer market size must satisfy Equation (57) 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:

$$W(K,T,x(.),p_2) = \int_0^T \Big(\frac{1}{t}u(x(t)) - \mathfrak{p}_1(t) - sx(t)\Big) f(t)\,\mathrm{d}t - c\Big(\frac{Y}{K}\Big)F(T) + \int_0^T \mathfrak{p}_1(t)f(t)\,\mathrm{d}t$$
$$+ (p_2 - f_1)Y - f_0(K) + (s - p_2)Y - JF \tag{61}$$

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 Equation (61) may be rewritten as

$$\begin{aligned} W(K,T,x(.),p_2) &= \int_0^T \Big(\frac{1}{t}u(x(t)) - sx(t)\Big) f(t)\,\mathrm{d}t - c\Big(\frac{Y}{K}\Big)F(T) + (p_2 - f_1)Y - f_0(K) + J\pi_C(J) \\ &= \int_0^T \Big(\frac{1}{t}u(x(t)) - sx(t)\Big) f(t)\,\mathrm{d}t - c\Big(\frac{Y}{K}\Big)F(T) + (p_2 - f_1)Y - f_0(K) \\ &= \int_0^T \frac{1}{t}u(x(t))f(t)\,\mathrm{d}t - c\Big(\frac{Y}{K}\Big)F(T) - f_1Y - f_0(K) - (s - p_2)Y \end{aligned} \tag{62}$$

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 Equations (59), (57), (55) and (60) 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

$$\frac{\partial W(K,T,x(.),p_2)}{\partial K} = \frac{F(T)}{K}c'\Big(\frac{Y}{K}\Big)\frac{Y}{K} - f_0'(K) \tag{63}$$

$$\frac{\partial W(K^e,T^e,x^e(.),p_2^e)}{\partial K} = 0 \tag{64}$$

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 Equation (62) as

$$W(K, T, x(.), p_2) = \int_0^T \Big(\frac{1}{t}u(x(t)) + (p_2 - s - f_1)x(t) - c\Big(\frac{X}{K}\Big)f(t)\,\mathrm{d}t\Big) - f_0(K) \qquad (65)$$

We can then find that

$$\frac{\partial W(K, T, x(.), p_2)}{\partial T} = \Big[\frac{1}{T}u(x(T)) + (p_2 - s - f_1)x(T) - c\Big(\frac{X}{K}\Big)\Big]f(T) \qquad (66)$$

$$\frac{\partial W(K^e, T^e, x^e(.), p_2^e)}{\partial T} = \frac{F(T^e)}{(T^e)^2}u(x(T^e)) > 0 \qquad (67)$$

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

Similarly, using Equation (65) to find the partial derivative of welfare with respect to the allocation $x(t)$,

$$\frac{\partial W(K, T, x(.), p_2)}{\partial x(t)} = \int_0^T \Big(\Big(\frac{1}{t}u'(x(t)) + (p_2 - s - f_1)\Big)f(t)\,\mathrm{d}t\Big) \qquad (68)$$

$$\frac{\partial W(K^e, T^e, x^e(.), p_2^e)}{\partial x(t)} = \frac{F(t)}{t^2}u'(x^e(t)) > 0 \qquad (69)$$

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 Equation (62)

$$\frac{\partial W(K, T, x(.), p_2)}{\partial p_2} = \frac{\partial Y}{\partial p_2}\Big(p_2 - f_1 - \frac{F(T)}{K}c'\Big(\frac{Y}{K}\Big)\Big) - \Big(s + Y\frac{\partial s}{\partial Y}\Big)\frac{\partial Y}{\partial p_2} + Y \qquad (70)$$

Using the first order condition for equilibrium termination fee, that is, Equation (60) , we have

$$\begin{aligned}\frac{\partial W(K^e, T^e, x^e(.), p_2^e)}{\partial p_2} &= -\Big(s^e + Y^e\frac{\partial s(Y^e)}{\partial Y}\Big)\frac{\partial Y(\mathfrak{p}_1^e(.), p_2^e, T^e)}{\partial p_2} \\ &= \frac{s^e Y^e}{p_2^e}\eta^e(Y, p_2)\Big(1 - \frac{1}{\eta^e(Y, s)}\Big) \qquad (71)\end{aligned}$$

where $\eta^e(Y, p_2)$ is the elasticity of content supply with respect to the termination fee at the equilibrium and $\eta^e(Y, s)$ is the elasticity of content demand with respect to the content usage price at the equilibrium.

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 too high relative to the socially optimal level if and only if $\eta^e(Y, s) < 1$.

Theorem 4(iv) says that whether the equilibrium termination fee is too high or too low with respect to the welfare-optimal levels depends on the elasticity of content demand with respect to the content usage price at the equilibrium. This may sound terse but the intuitive reason behind this result becomes clear by reasoning through the model. What happens to welfare as a result of a small increase in p_2 at the equilibrium can be found by answering what happens to the aggregate outflow from consumers in terms of content usage payments i.e. by figuring out the sign of $\frac{\partial(-s(Y)Y)}{\partial p_2}$ at the equilibrium. The first order effect is positive. As content flow on the network goes down as a result of a small increase in p_2, the aggregate outflow goes down as well. The second order effect, however, is negative. The lower supply of content drives up the price in the content market and contributes to a greater payment outflow from consumers. If content price is sufficiently responsive to changes in demand or equivalently if the elasticity of demand with respect to the content price is sufficiently low, the negative second order effect dominates the positive first order effect to drive down welfare as a result of a small increase in p_2.

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 Equation (70), we have

$$\begin{aligned}\frac{\partial W}{\partial p_2} &= p_2\frac{\partial Y}{\partial p_2} + Y - \frac{\partial Y}{\partial p_2}\Big(e + s + Y\frac{\partial s}{\partial Y}\Big) \\ &= p_2\frac{\partial Y}{\partial p_2} + Y\Big[1 + \frac{\eta(Y,p_2)}{p_2}\Big(e + s\big(1 - \frac{1}{\eta(Y,s)}\big)\Big)\Big] \end{aligned} \tag{72}$$

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 price. Then a sufficient condition for $\frac{\partial W(p_2=0)}{\partial p_2} < 0$ is that $e^0 + s^0\big(1 - \frac{1}{\eta^0(Y,s)}\big) < 0$ 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 imposing a positive termination fee reduces welfare. On the other hand, if content flow is elastic with respect to the content price, then a 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. 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.

Acknowledgements

All the authors declare that they have no relevant material or financial interests related to this research.

Appendix

References

ARMSTRONG, M. (2006): "Competition in two-sided markets," *The RAND Journal of Economics*, 37, 668–691.

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.*

ECONOMIDES, N. AND B. E. HERMALIN (2012): "The economics of network neutrality," *The RAND Journal of Economics*, 43, 602–629.

ECONOMIDES, N. AND J. TÅG (2012): "Network neutrality on the Internet: A two-sided market analysis," *Information Economics and Policy*, 24, 91–104.

FARRELL, J. AND G. SALONER (1985): "Standardization, compatibility, and innovation," *the RAND Journal of Economics*, 70–83.

——— (1986): "Installed base and compatibility: Innovation, product preannouncements, and predation," *The American economic review*, 940–955.

GREENSTEIN, S. (2020): "The Basic Economics of Internet Infrastructure," *Journal of Economic Perspectives*, 34, 192–214.

GREENSTEIN, S., M. PEITZ, AND T. VALLETTI (2016): "Net neutrality: A fast lane to understanding the trade-offs," *Journal of Economic Perspectives*, 30, 127–50.

HEMPHILL, C. S. (2008): "Network neutrality and the false promise of zero-price regulation," *Yale J. on Reg.*, 25, 135.

KATZ, M. L. AND C. SHAPIRO (1985): "Network externalities, competition, and compatibility," *The American economic review*, 75, 424–440.

——— (1986): "Technology adoption in the presence of network externalities," *Journal of political economy*, 94, 822–841.

LEE, R. S. AND T. WU (2009): "Subsidizing creativity through network design: Zero-pricing and net neutrality," *Journal of Economic Perspectives*, 23, 61–76.

MACKIE-MASON, J. K. AND H. R. VARIAN (1995): "Pricing congestible network resources," *IEEE journal on Selected Areas in Communications*, 13, 1141–1149.

Maskin, E. and J. Riley (1984): "Monopoly with incomplete information," *The RAND Journal of Economics*, 15, 171–196.

McManus, M. (1962): "Numbers and size in Cournot oligopoly," *Bulletin of Economic Research*, 14, 14–22.

——— (1964): "Equilibrium, numbers and size in Cournot oligopoly," *Bulletin of Economic Research*, 16, 68–75.

Mussa, M. and S. Rosen (1978): "Monopoly and product quality," *Journal of Economic theory*, 18, 301–317.

Oi, W. Y. (1971): "A Disneyland dilemma: Two-part tariffs for a Mickey Mouse monopoly," *The Quarterly Journal of Economics*, 85, 77–96.

Pil Choi, J. and B.-C. Kim (2010): "Net neutrality and investment incentives," *The RAND Journal of Economics*, 41, 446–471.

Roberts, J. and H. Sonnenschein (1976): "On the existence of Cournot equilbrium without concave profit functions," *Journal of Economic Theory*, 13, 112–117.

Rochet, J.-C. and J. Tirole (2003): "Platform competition in two-sided markets," *Journal of the european economic association*, 1, 990–1029.

——— (2006): "Two-sided markets: a progress report," *The RAND journal of economics*, 37, 645–667.

Schmalense, R. (1981): "Monopolistic two-part pricing arrangements," *The Bell Journal of Economics*, 445–466.

Shapiro, C. and H. R. Varian (1998): *Information rules: a strategic guide to the network economy*, Harvard Business Press.

Varian, H. R. (1989): "Price discrimination," *Handbook of industrial organization*, 1, 597–654.

Wu, T. (2003): "Network neutrality, broadband discrimination," *J. on Telecomm. & High Tech. L.*, 2, 141.