Pricing, Competition and Content for Internet Service Providers

Peter Key, *Fellow, IEEE,* and Richard Steinberg, *Member, IEEE*

***Abstract*—We examine competition between two Internet Service Providers (ISPs), where the first ISP provides basic Internet service, while the second ISP provides Internet service plus content, i.e., *enhanced service*, where the first ISP can partner with a Content Provider to provide the same content as the second ISP. When such a partnering arrangement occurs, the Content Provider pays the first ISP a transfer price for delivering the content. Users have heterogeneous preferences, and each in general faces three options: (1) buy basic Internet service from the first ISP; (2) buy enhanced service from the second ISP; or (3) buy enhanced service jointly from the first ISP and the Content Provider. We derive results on the existence and uniqueness of a Nash equilibrium, and provide closed-form expressions for the prices, user masses, and profits of the two ISPs and the Content Provider. When the first ISP has the ability to choose the transfer price, then when congestion is linear in the load, it is never optimal for the first ISP to set a negative transfer price in the hope of attracting more revenue from additional customers desiring enhanced service. Conversely, when congestion is sufficiently super-linear, the optimal strategy for the first ISP is either to set a negative transfer price (subsidizing the Content Provider) or to set a high transfer price that shuts the Content Provider out of the market.**

***Index Terms*—communication networks, competition, content provider, optimal pricing, Nash equilibrium, profit**

I. Introduction

Several years ago, Comcast complained to the FCC that Netflix was asking for special access to its broadband network. Comcast said that the issue could cause a financial dispute, but it did not require the involvement of regulators. Netflix's response to the FCC was that it was not seeking special treatment, and was being pressured by large operators having market power to pay for improved delivery of its content [1].

In this paper we address questions of pricing and competition in a market where users connect to a service provider that either offers basic Internet service or *enhanced service*, i.e., Internet service plus content. This could be low bit-rate stream content, such as audio and certain types of video, or high bit-rate stream content, such as high-definition television.

We characterize the enhanced service by the *relative bandwidth*, b, it requires as compared with basic service. If the relative bandwidth $b=1$, this means that the enhanced service takes essentially the same bandwidth as the basic service, i.e., the content requires negligible additional bandwidth; for example, when the service consists of streaming music over a well-provisioned fixed network. If the bandwidth $b>1$, this means that the enhanced service places an extra load on the network. In cases where the relative bandwidth b is large, this implies that the content requires high bandwidth compared to basic Internet usage; for example, delivering movies over the existing infrastructure. (Relative bandwidth b is defined precisely in Section III.)

In our model we consider the existence of an ISP that offers basic Internet service (network 1) competing with a second ISP (network 2) that offers Internet service plus content. In addition, we assume the existence of a third-party Content Provider (CP) that offers the same content as is provided by network 2, and which can partner with network 1 to jointly provide the content to consumers, where network 1 charges the Content Provider a *transfer price* for delivering the content. We make no restrictions on the transfer price; in particular, it can be positive, negative—i.e., a subsidy, perhaps imposed by a regulator—or zero.

Users will generally differ in their *willingness to pay*, w, for content. Each user has, in general, three options: (Option 1) buy basic Internet service from network 1 at price p_1, (Option 2) buy enhanced service from network 2 at price p_2, or (Option 3) buy enhanced service jointly from network 1 and the Content Provider at price p_1+p_3. In choosing which network to join, a user takes into consideration not only the type of service and total price he would need to pay, but also the level of congestion on the network.

As an example of this scenario, consider an Internet Service Provider (network 1) partnering with Netflix (CP), where Sky Broadband (network 2) is also available to users. Netflix offers television programs plus movies, but no basic Internet service; thus, users need to make arrangement with the Internet Service Provider in order to access the Netflix content. Sky Broadband provides a package of basic Internet service plus content, where the content is essentially the same as that provided by Netflix. The Internet Service Provider offers a choice to each user of either: (Option 1) basic Internet service at around p_1 = \$40 per month; or (Option 3) basic Internet service plus content from Netflix, at $p_1 + p_3$ = \$50 per month, \$40 of which goes to the Internet Service provider (p_1) and \$10 goes to Netflix (p_3). Sky Broadband offers basic Internet service together with content, at around p_2 = \$50 per month (option 2). The ISP currently charges Netflix a proprietary per-subscriber monthly transfer fee, but is considering dropping it, or even

Manuscript received ...

Peter Key was formerly with Microsoft Research, Cambridge, CB1 2FB, UK, email:peter.key@outlook.com

Richard Steinberg is with the London School of Economics, Department of Management, NAB 3.08, Houghton Street, London WC2A 2AE, UK, email:r.steinberg@lse.ac.uk. He would like to thank the Statistical Laboratory, University of Cambridge where he was a visitor during his sabbatical year 2018-19, during which time this paper was completed.

This article has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the authors.

replacing it with a subsidy, in order to increase its own network traffic.

We consider the following questions: What will be the prices charged to users in equilibrium by the basic service ISP, the ISP offering its own content, and the Content Provider? Under what conditions will the Nash equilibrium be unique? In a scenario where network 1 can choose the transfer price as leader in a Stackelberg game with the objective of maximizing its own revenue, will network 1 ever find it optimal to set a negative transfer price in the hope of attracting more revenue from the additional customers desiring enhanced service? Will network 1 ever decide to not charge a transfer price?

Not charging a transfer price has an interpretation in terms of net neutral pricing. It is interesting to note that while the initial meaning of a "net neutral" traffic regime was one that does not discriminate in terms of *transmission* based on packet content, the debate has enlarged to the point where the phrase can now refer to a regime that does not discriminate in terms of *price* based on packet content [2]. In late 2016, Comcast and Netflix announced a tie-up to allow Comcast customers to access Netflix services through Comcast's set-top box, causing some to argue that the set-top box is the new frontier for net neutrality [3]. Economides and Hermalin [2] stress the importance of competition in the context of net-neutral pricing, and emphasize that the price a customer pays to an ISP for Internet access depends crucially on the availability of competing ISPs for this customer.

Our results are as follows. When the transfer price is fixed and known *ex ante* by all three firms, if a Nash equilibrium in positive prices exists with users on each of the three services, then the equilibrium will be unique, and we provide closed-form expressions for the prices, user masses, and profits of the two networks and the Content Provider. We also provide necessary and sufficient conditions for the prices to constitute a Nash equilibrium in which networks 1 and 2 and the Content Provider all have users. Further, we characterize the degenerate equilibria and provide closed form expressions for them.

The shape of the congestion function affects the impact of the transfer price in the Stackelberg game. First, we consider the case of linear congestion, reflecting a situation where both network providers have well-provisioned networks. We show that in this case it is never optimal for network 1 to set a negative transfer price in the hope of attracting more revenue from the additional customers desiring enhanced service. Neither is it ever optimal for network 1 to set a zero transfer price. The optimal transfer price uniquely determines one of two outcomes: either it enables all three parties to make a profit, or it shuts out the Content Provider, thereby creating a duopoly. We characterize the transfer price associated with each of these two outcomes, and also describe necessary and sufficient conditions for each outcome to occur.

Second, we consider the case of non-linear congestion, reflecting situations where the networks are not well provisioned or where demand exceeds supply. We show that, as soon as congestion becomes sufficiently non-linear (e.g., quadratic), if network 1 does not shut out the Content Provider with a high price, then network 1's optimal strategy is to subsidize the Content Provider (i.e., set a negative transfer price) to capture extra revenue for the additional customers desiring enhanced service.

The remainder of the paper is organized as follows. In Section II, we discuss the related literature. In Section III, we present the base model. In Section IV, we consider optimizing the transfer price in a Stackelberg game. In Section V, we relax the assumptions of the base model, by allowing the sensitivity to congestion to be service dependent, and by considering non-linear congestion. We conclude in Section VI. In the interest of concision, almost all proofs in the text are sketch proofs, where the complete proofs are given in the Supplementary Material. The only exceptions are Theorem III.16 where the proof is short and is given in its entirety in the text, and Theorem IV.2, which is a compilation of several results proved earlier in the paper.

II. Related Literature

This paper is related to both the literature on charging schemes for congestible resources, and the literature on congestion games in communication networks. From the charging scheme literature, the most relevant paper to our work is de Palma and Leruth [4], which examined duopoly outcomes for two firms in a setting similar to our base model. However, [4] models two firms offering an identical service, in contrast to our model in which one firm offers a basic service and the other firm offers an enhanced service, but the first firm can offer the enhanced service by partnering with a third firm.

The literature on congestion games in communication networks is large and growing rapidly. Marden and Wierman [5] considered a game-theoretic approach to the study of utility design for distributed resource allocation. They introduced a class of games they refer to as "distributed welfare games," and demonstrate that cost sharing methodologies are beneficial for utility design. Their work has a broad range of applications that includes communication networks.

Gibbens, Mason and Steinberg [6] considered competition between two networks, each of which may offer multiple services classes generated by subdivision of the network into subnetworks, differentiated only by capacity, price, and the consequent level of congestion, i.e., "Paris Metro pricing." Specifically, their model has two competing, profit-maximizing Internet Service Providers, each of which may offer either one or two service classes. In the case where an ISP chooses to offer two service classes, it forms them by logically dividing its network in two, and charging separate prices on each subnetwork. Congestion on a subnetwork is determined in equilibrium by the fraction of the first ISP's total network capacity allocated to a subnetwork, and the number of users on the subnetwork. The main result is that, in the unique equilibrium outcome, neither ISP subdivides its network and the two firms charge the same price. These results tend to indicate that Paris Metro Pricing will not be viable in a competitive market.

Shetty, Schwartz and Walrand [7] considered a scenario of multiple identical competing ISPs, where each provides the same two services. The main focus is on capacity investment, which occurs in the first stage of the model, where the

ISPs simultaneously and independently invest in irreversible capacity and observe the total capacity among all the ISPs. In the second stage, each firm allocates its capacity to the services. In the third stage of the model, the ISPs make pricing decisions, where the access price of each service is identical for all the ISPs, which is taken to be the minimum price among the ISPs.

Johari, Weintraub and Roy [8] studied oligopolistic competition in service industries with congestion effects. They showed that, in settings that exhibit non-increasing returns to investment, if a pure-strategy Nash equilibrium exists, then it is unique. Their paper also provides conditions for the existence of a pure-strategy Nash equilibrium. Finally, they extend their model to one in which providers can decide whether to enter the market. The authors find that the equilibrium number of entrants exceeds the socially efficient level, but that entry becomes efficient asymptotically as the sunk entry cost becomes small. They conclude by stating that their work leaves many significant directions for future research. The first direction they discuss arises from their assumption that consumers have homogeneous preferences. The write: "We leave for future research study of a model where consumers have heterogeneous preferences."

In our model, we consider a form of oligopolistic competition with congestion effects, where consumers have heterogeneous (dis)preferences for congestion. We provide conditions for the existence and uniqueness of a pure-strategy Nash equilibrium, as well as closed-form expressions for the prices, user masses, and profits.

Ma and Misra [9] look at related issues, although their setting is different than ours. They consider consumer utility as their primary objective, rather than provider profit. In their model, ISPs offer two classes of service to Content Providers, ordinary and premium, where CPs get charged extra for sending traffic in the premium class. The authors consider both monopolistic and oligopolistic scenarios; in their oligopolistic scenario they allow a different Quality of Service mechanism for premium, whereas we use sharing at the packet or bit level. They have elastic demand, dependent on capacity provided for the service; in contrast, we have user demand directly affected by prices. In their scenario, they find allocations being proportional to the capacity of the Internet Service Providers.

In a more recent paper, Ma [10] looks at subsidizing competition among Content Providers. He models a single (access) ISP serving users who access a set of service providers, with different groups of users associated with each CP, where each CP can subsidize its users by a fixed amount. Ma first assumes the CPs bid strategically and the ISP's price is fixed. He shows that a Nash equilibrium exists under mild conditions, and that subsidies are either the maximum allowable subsidy, or a function of the profit margin of the CP and its elasticity metrics. He then considers a sequential game in which the regulator sets the maximum allowable subsidy, the ISP sets the price, and the CPs react. In this scenario he finds that allowing more competition will increase welfare, but also will provide the ISP with an incentive to raise prices, discouraging user demand, subsidies, and throughput. In contrast, we consider competition among ISPs and a CP, where one ISP can be considered both an ISP and a CP. Unlike Ma's model, the users pay both the CP for the enhanced service and the ISP for carrying the traffic. We allow the ISPs to charge an additional transfer price to a CP: a positive transfer price passes money from the CP to the ISP. This is like a subsidy in Ma's model but, unlike Ma's model, the transfer price doesn't affect the price users pay. In contrast to Ma, we also allow negative transfer prices, whereby the ISP subsidizes the CP, like a price rebate. A distinctive feature of our approach is that we show the critical role played by network congestion in determining a firm's behavior. Specifically, we find that when congestion is linear, the optimal strategy for the ISP is to set a non-negative transfer price, i.e., extract money from the CP—equivalent to a CP in Ma's model subsiding users—or to price the CP out of the market. However, when congestion is superlinear, it can be beneficial for the ISP to subsidize the CP to encourage traffic.

III. The Model

We model a setting where three firms compete to maximize individual profits: Network 1 provides basic service, network 2 provides enhanced service, and a Content Provider provides enhanced service over network 1. As discussed in [6], there is good reason to suppose that, under certain circumstances, industries with congestion may have very concentrated market structure; that is, there exist a small number of firms collectively having a large market share. (See for example [11], [12].) Further, as Gibbens et al. [6] point out, this setting is the most transparent environment in which to study the effect of competition on the use of multiple service classes.

We assume that a user pays a price *per unit time* for the right to be connected to and receive service or services from network i. Thus, network prices are subscription-based. In [13], Cachon and Feldman consider the question, "How should a firm price its service when congestion is an unavoidable reality?" They point out that some firms sell subscriptions for their service, citing as an example the Internet Service Provider AOL. AOL initially charged customers per-use access fees, but later switched to subscription pricing in the form of a monthly fee with no usage limitation. The authors find that subscription pricing is more effective at earning revenue than per-use access fees. They conclude that subscription pricing can be effective even if congestion is relevant for the overall quality of the service.

A. The users

On joining network i, a user w receives quasi-linear utility $U(w;i)$ per unit time, which has four components: (i) a positive benefit V of receiving Internet service; (ii) a dis-benefit depending on the degree of congestion on the network k_i, scaled by the sensitivity to congestion g; (iii) a dis-benefit from having to pay a price $p_i \geq 0$ to network i; and, if network i is providing enhanced service, (iv) a positive benefit w which is the consumer willingness to pay for the content.

A user generally has three options: (Option 1) purchase *basic service* from network 1 at price p_1, (Option 2) purchase *enhanced service* from network 2 at price p_2, and (Option 3) purchase *enhanced service* from network 1 and the Content

Provider, consisting of basic service from network 1 at p_1 and content from the Content Provider at price p_3. However, it is possible that the third option is unavailable. (See §*B*, below.)

We will assume that the user sensitivity to congestion is the same for both basic and enhanced services, i.e., $g^E = g^B \equiv g$, and that congestion is a function of the load, ρ on the network, defined below. We refer to the quantity

$$K_i := K_i(g, \rho_i) = k(g\rho_i) \tag{1}$$

as the *scaled congestion* on the network i where ρ_i is the load on network i, and k a non-negative function.

The utility to a user w choosing Option 1, 2, or 3, respectively, is:

$$U(w, B; 1) \;=\; V - K_1 - p_1 \tag{2}$$

$$U(w, E; 2) \;=\; V - K_2 - p_2 + w \tag{3}$$

$$U(w, E; 3) \;=\; V - K_1 - p_1 - p_3 + w. \tag{4}$$

We assume that each user joins one and only one network, i.e., the covered market assumption. In our scenario, this means that, for given parameter values, the benefit of Internet service, V, is sufficiently large relative to the congestion cost and price.

Users differ in their willingness to pay for enhanced service. Those receiving little utility from enhanced service will have low values of w, and those receiving great utility will have high values of w. We assume that there is a continuum of users whose w parameters form a population distribution that is given by a uniform probability distribution with support $[0, w_{max}]$. (See Gibbens et al. (2000) for a discussion of the choice of the uniform distribution in this context.) We normalize by taking the support to be $[0, 1]$, under the mapping $V \mapsto V/w_{max}$, $g \mapsto g/w_{max}$, $p_i \mapsto p_i/w_{max}$, $w \mapsto w/w_{max}$. Thus, $0 \leq w \leq 1$.

B. The firms

Costs to each firm of providing their service are assumed sunk, and thus for simplicity all costs are set to zero. We allow network 1 to charge the Content Provider a *transfer price* t. For now, in this base model, we assume that the transfer price is fixed and known *ex ante* by all three firms. This could occur, for example, in the situation where the transfer price is set by an outside agency such as a regulator. (Later, in Section IV, we allow network 1 to optimize the price.) Specifically, the Content Provider will be obliged to pay network 1 the sum tQ_{13} for delivering the content for a user mass of size Q_{13}. A negative transfer price is possible, in which case network 1 is paying a *subsidy* $s := -t$ to the Content Provider in order to carry its service. A zero transfer price is possible. It is also possible that network 1 chooses to set a sufficiently high transfer price that the Content Provider is priced out of the market. In such a case, the users' third option is eliminated, viz., users will not have the option of purchasing enhanced service from network 1 together with the Content Provider, although they will still have the option of purchasing enhanced service from network 2.

It is worth stressing that, throughout this paper, the extra service that customers are prepared to pay for is *content*, and not simply an improvement in Quality of Service for basic Internet service. Moreover, we assume that the quality of service seen by the differing services offered is identical when they are carried on the same network, and hence there is non-discrimination at the service level.

Let Q_{11} denote the *mass* of users that buys basic service from network 1; let Q_2 denote the mass of users that buys enhanced service from network 2; and let Q_{13} denote the mass of users that buys enhanced service jointly from network 1 and the Content Provider. Then $Q_1 = Q_{11} + Q_{13}$ is the mass of users making use of network 1.

Each user of basic service over a network requires expected bandwidth b^B, while each user of enhanced service over a network requires expected bandwidth b^E, with $b^E \geq b^B$. Further, *congestion* on network i is a convex function of the *load*, ρ_i, on the network. defined to be the sum of the expected bandwidth per user on the network, b^x, times the mass of users on the network, all divided by the *capacity* of the network, C_i, i.e.

$$\rho_1 \;=\; \frac{b^B Q_{11} + b^E Q_{13}}{C_1}, \qquad \rho_2 \;=\; \frac{b^E Q_2}{C_2}. \tag{5}$$

For now we assume the function $k(\cdot)$ is linear and, without loss of generality, we assume $k(x) = x$. From (1) and (5), the expressions for the K_i are:

$$K_1 \;=\; g\frac{b^B Q_{11} + b^E Q_{13}}{C_1}, \qquad K_2 \;=\; g\frac{b^E Q_2}{C_2}. \tag{6}$$

With these assumptions, the profit generated for network i is given by

$$\pi_1 \;=\; p_1 Q_1 + tQ_{13}, \;\; \pi_2 \;=\; p_2 Q_2, \;\; \pi_3 \;=\; (p_3 - t)Q_{13}. \tag{7}$$

To ease exposition and without loss of generality, we shall normalize masses and capacities so that the size of the market, i.e., the total mass Q, is 1. Therefore the Q_i represent the fraction of the market that firm i has captured. This is equivalent to scaling both Q_i and C_i by $1/N$, where N is the number of users in the market. Under this scaling, (where $Q = 1$), the profit π represents the per-user profit, and needs to be scaled back by N to recover the total profit. We have scaled C_i, Q_i so that the total mass of users (Q) is 1, hence $Q_1 + Q_2 = 1$. Note that by virtue of our previous normalization of w, we now have a unit mass of users whose w parameters are uniformly distributed on $[0, 1]$. (There is no loss in generality by normalizing on both the willingness-to-pay and the size of the market, due to the fact that these are independent parameters of the problem.)

We define the *effective capacities*, $\widehat{C}_1$ and $\widehat{C}_2$, which are dimensionless quantities, as follows:

$$\widehat{C}_1 \;:=\; \frac{C_1}{b^B g}, \qquad \widehat{C}_2 \;:=\; \frac{C_2}{b^E g}. \tag{8}$$

The effective capacity of a network can be interpreted as the mass of users that network i can tolerate before "saturating," taking into account not only bandwidth dependent on service provided, but also the important factor of sensitivity to price.

We define the *effective capacity ratio*, r, as:

$$r \;:=\; \widehat{C}_2/\widehat{C}_1. \tag{9}$$

The utilities, and hence the solution, will depend on the $\widehat{C}_i$ and r, and also on the *relative bandwidth*:

$$b \ := \ b^E/b^B. \tag{10}$$

Substituting for the K_i from (6), and using the definitions of $\widehat{C}_i$, r, and b from (8), (9) (10), allows us to write:

$$K_2 - K_1 \ = \ \frac{1}{r\widehat{C}_1}[Q_2 - r(Q_{11} + bQ_{13})]. \tag{11}$$

This expression will enable us to eliminate the congestion terms in most of the expressions we derive in the remainder of the paper.

C. User choices

We assume that users are utility maximizers, and hence choose the option (1, 2, or 3) that is individually best for them. Users who are more willing to pay for content will be more willing to choose network 2 over network 1, depending on the differences in scaled congestion and in the prices on the two networks. Similarly, users who are more willing to pay for content will be more willing to choose option 3 over network 1; specifically, a user will choose option 3 over network 1 if their willingness to pay for content exceeds the Content Provider's price. In other words, there are two critical values, denoted ∇_{12} and ∇_{13}, whereby users with a willingness to pay of at least ∇_{12} weakly prefer network 2 to network 1, and users with a willingness to pay of at least ∇_{13} weakly prefer option 3 to network 1. These critical values are, from (2), (3), (11), and from (2), (4), respectively:

$$\begin{aligned} \nabla_{12} \ &= \ K_2 - K_1 + p_2 - p_1 \\ &= \ \frac{1}{r\widehat{C}_1}[Q_2 - r(Q_{11} + bQ_{13})] + p_2 - p_1 \qquad (12) \\ \nabla_{13} \ &= \ p_3. \qquad (13) \end{aligned}$$

By comparing user preferences, it is straightforward to show:

Lemma III.1. *In equilibrium:*

1) *If* $[Q_2 - r(Q_{11} + bQ_{13})]/r\widehat{C}_1 < p_1 + p_3 - p_2$ *(equivalently,* $\nabla_{12} < \nabla_{13}$*), each user chooses network 1 or network 2, according to whether* w *is, respectively, smaller or larger than* ∇_{12}.
2) *If* $[Q_2 - r(Q_{11} + bQ_{13})]/r\widehat{C}_1 > p_1 + p_3 - p_2$ ($\nabla_{12} > \nabla_{13}$), *each user chooses network 1 or option 3, according to whether* w *is, respectively, smaller or larger than* ∇_{13}.
3) *If* $[Q_2 - r(Q_{11} + bQ_{13})]/r\widehat{C}_1 = p_1 + p_3 - p_2$ ($\nabla_{12} = \nabla_{13}$), *each user with* $w < \nabla_{12} = \nabla_{13}$ *chooses network 1. Users with* $w > \nabla_{12} = \nabla_{13}$ *split between network 2 and option 3.*

D. Prices

A set of prices $\mathbf{p} = \{p_i\}$, $i = 1, 2, 3$ is *feasible* if the user masses $\mathbf{Q} = \{Q_{11}, Q_{13}, Q_2\}$ are non-negative and together cover the market, and if the profits of the firms π_i, $i = 1, 2, 3$ are non-negative. I.e.,

$$\begin{aligned} Q_{11} &\geq 0 \qquad (14) \\ Q_{13} &\geq 0 \qquad (15) \\ Q_2 &\geq 0 \qquad (16) \\ Q_{11} + Q_{13} + Q_2 &= 1 \qquad (17) \\ \pi_i &\geq 0 \quad i = 1, 2, 3 \qquad (18) \end{aligned}$$

We now show that any non-negative set of prices will result in user masses that are non-negative and cover the market, and will partition the non-negative orthant $\{p_i \geq 0\}$ in price space. Furthermore, for a given transfer price t, *feasibility* partitions the non-negative orthant into three regions of feasible prices $\mathbf{p}$ corresponding to the three cases of Lemma III.1:

Case 1: Region 1, where $\nabla_{\mathbf{12}} < \nabla_{\mathbf{13}}$.
Users split between Options 1 and 2, i.e., $Q_{13} = 0$.

Case 2: Region 2, where $\nabla_{\mathbf{12}} > \nabla_{\mathbf{13}}$.
Users split between Options 1 and 3, i.e., $Q_2 = 0$.

Case 3: Region 3, where $\nabla_{\mathbf{12}} = \nabla_{\mathbf{13}}$.
Users indifferent between Options 2 and 3.

Throughout the paper we use the terminology "Case 1" as defined above, and similarly for Cases 2 and 3.

Note that for a given transfer price t, there can be "holes" in feasible $\mathbf{p}$ space that represent infeasible prices (where the allocations are non-negative but one or more of the firms' profits are negative).

The demand allocations are determined by users selfishly seeking to maximize their own utility, i.e., a Wardrop equilibrium [14] for a continuum of users where user w chooses option i if $U(w,i) \geq U(w,j) \ \forall j$. Existence of such an equilibrium in our setting follows from the standard results.

Theorem III.2. Existence and Uniqueness of User Masses. *Given any feasible vector of prices* $\mathbf{p}$ *and any transfer price* t*, there exists a unique vector of user masses* $\mathbf{Q}$.

Note that it is the user masses Q_i that are uniquely determined, which are given respectively by the Lebesgue measure of the corresponding user choice sets, rather than the individual user choices.

Corollary III.3. *When* $t \geq 0$*, the feasible region in 3 dimensional price space* p_i *is the union of three polyhedrons, each of which corresponds to one of the three cases of Lemma III.1. Each volume is formed by the intersection of half-planes generated by hyperplanes, each hyperplane corresponding to one of the conditions* (14) *to* (18)*. When* $t < 0$*, a similar result holds, except the region corresponding to Case 3 of Lemma III.1 is convex but not in general polyhedral. For all* t*, the feasible support for each* $\pi_i(p_i) : \mathbb{R} \mapsto \mathbb{R}$ *is a convex interval on the line.*

Sketch Proof. In each of the three cases of Lemma III.1, $\mathbf{Q}$ is an affine function of $\mathbf{p}$. Hence in all cases, each constraint on the users masses, (14) to (17), corresponds to a separating hyperplane in $\mathbf{p}$-space. In the case $t = 0$, each of the profit constraints (18) corresponds to a separating hyperplane or the space formed by intersecting hyperplanes (e.g., for $\pi_2 \geq 0$ is equivalent to $p_2 \geq 0$ and $Q_2 \geq 0$). The same holds true when $t \geq 0$, apart from the constraint for $\pi_1 \geq 0$ for Case 3, which is quadratic in p_1 but which reduces to an intersection of hyperplanes. When $t < 0$, it is straightforward to show the region is convex. Combining these statements proves the corollary.

□

E. Nash equilibria

Definition III.4. *(Nash equilibrium) Given a transfer price t, a vector of prices $p^*(t) = (p_1^*(t), p_2^*(t), p_3^*(t))$ is a Nash equilibrium if for all $p_i(t)$ ($i = 1, 2, 3$),*

$$\pi_i(p_i^*(t); p_{-i}^*(t)) \geq \pi_i(p_i(t); p_{-i}^*(t)) \tag{19}$$

where $p_{-i}^*(t)$ is the vector of prices $p^*(t)$ excluding $p_i^*(t)$.

Thus for a given transfer price t, in a Nash equilibrium no firm can increase its own profits by unilaterally changing its price p_i. Note this assumes that the networks employ pure strategies. As pointed out by Gibbens et al. (2000), a standard criticism of mixed strategy equilibria is that they impose too large an informational burden on users. In choosing a network to join, users are faced with price distributions from which the final prices will be drawn, rather than specified price levels. Moreover, in order for users to decide which network to join, they must be aware not only of the equilibrium strategies of the networks, which are probability distributions over prices, but also of the choices of all other users. It is seems highly unlikely that actual users would be capable of performing this task. See also chapters 8 and 10 of [15]. A strategy for a user is a choice of network to join, given the prices quoted by the network. We now characterize the Nash equilibria.

Theorem III.5. Uniqueness of the Nash Equilibrium. *If a Nash equilibrium exists with positive prices $\{p_i^*\}$, with networks 1 and 2 and the Content Provider each having users and making positive profit, then there will be a unique equilibrium:*

$$p_1^* = \frac{(2+br)[2/r+1+b+\widehat{C}_1] + [(b-1-5\widehat{C}_1)(1+r) - 4\widehat{C}_1^2 r]\, t}{2\widehat{C}_1[3 + (2+b+2\widehat{C}_1)r]} \tag{20}$$

$$p_2^* = \frac{2/r + 2(1+2b+\widehat{C}_1) + b(3+b+3\widehat{C}_1)r - A\, t}{2\widehat{C}_1[3 + (2+b+2\widehat{C}_1)r]} \tag{21}$$

where $A := [(b-1+\widehat{C}_1)(1+r) - 2b\widehat{C}_1 r]$

$$p_3^* = \frac{(2+br) + [3 + r(3+4\widehat{C}_1)]\, t}{2[3 + (2+b+2\widehat{C}_1)r]} \tag{22}$$

with user masses:

$$\begin{aligned} Q_1^* &= \frac{r\widehat{C}_1}{1+br}(p_1^*+t), \quad Q_2^* = \frac{r\widehat{C}_1}{1+br}p_2^*, \\ Q_{13}^* &= \frac{1+r(1+\widehat{C}_1)}{1+br}(p_3^*-t) \end{aligned} \tag{23}$$

where the profits π_1, π_2, and π_3 can be calculated from (7).

Sketch Proof. Only in Case 3 of Lemma III.1 do network 1, network 2, and the Content Provider all have users. There, by direct calculation, it is straightforward to show profit functions $\pi_i(p_i)$ are strictly concave, Hence a local optimum, and thus a candidate Nash equilibrium, exists when the first order conditions are satisfied, resulting in prices p_i^* as stated. □

Technical Remark III.6. *The optimal allocation can also be expressed terms of the optimal prices alone. In the case $t = 0$,*

$$\begin{aligned} Q_1^* &= \frac{p_1^*}{p_1^*+p_2^*}, \quad Q_2^* = \frac{p_2^*}{p_1^*+p_2^*}, \\ Q_{13}^* &= \frac{p_1^*}{p_1^*+p_2^*} - p_3^* \end{aligned} \tag{24}$$

Theorem III.9 given below, shows that a Nash equilibrium will exist where all three firms have customers and make a profit, provided certain conditions are satisfied. The conditions can be expressed as bounds on t related to the parameters of the problem, but only in the $t=0$ case is there a simple characterization of sufficient conditions for the bounds to be met.

Example III.7. *We now consider the case where the networks have the same capacity, $C_2 = C_1$, and set $\widehat{C}_1 = 1$, hence $r = 1/b$. If t is negative, and hence a subsidy, then the amount of users on network 1 (taking only basic service) decreases compared to the case of no transfer price, while the mass of users availing themselves of the Content Provider increases, and the mass of users on network 2 increases slightly. When t is positive, the reverse is true. A subsidy (negative t) has the effect of decreasing the profit of network 1, while increasing the profits of network 2 and the Content Provider, while the converse is true for positive t. The percentage profit difference is most marked for the Content Provider.*

Figure 1 illustrates how the combined profit of network 1 and the Content Provider ($\pi_1^* + \pi_3^*$) compares with that of network 2 (π_2^*) as we vary both the relative bandwidth, b, and transfer price t, with $r = 1/b$. Here, b varies between 1 and 6, t varies between -0.5 and 0.5, and $\widehat{C}_1 = 1$. The case of no transfer-price corresponds to the line $t = 0$ in this figure. The effect of b can be clearly seen. Note that increasing b appears to provide larger profits. Indeed, in any setting where prices increase with congestion, there is an apparent incentive to decrease capacity and hence increase prices and thus profit. However, this ignores the fact that users will only be prepared to endure a certain level of congestion before opting out of the service entirely. The latter scenario corresponds to leading to a reduction in user mass. Thus, the graphs in this paper need to be interpreted primarily as illustrating behavior among the users conditioned on their parameters, rather than comparing results at different parameter values.

Definition III.8. (ϵ-equilibrium*) Given a transfer price t, a vector of prices $p^*(t) = (p_1^*(t), p_2^*(t), p_3^*(t))$ is an* ϵ-equilibrium *for $\epsilon \geq 0$ if, for all $p_i(t)$ ($i = 1, 2, 3$),*

$$\pi_i(p_i^*(t); p_{-i}^*(t)) \ \geq \ \pi_i(p_i(t); p_{-i}^*(t)) - \epsilon \tag{25}$$

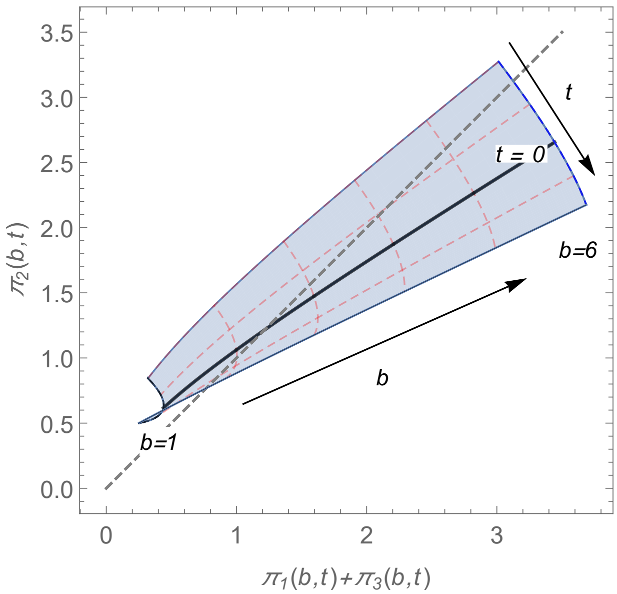

Fig. 1. $\pi_1 + \pi_3$ vs. π_2 as b and t vary. Solid lines show $b = 1$ and $t = 0$ boundaries.

where $p^*_{-i}(t)$ is the vector of prices $p^*(t)$ excluding $p^*_i(t)$.

Thus for a given transfer price t, in an ϵ-equilibrium no firm can increase its own profits by more than ϵ by unilaterally changing its price p_i. (See [16] and [17].)

The first part of the following theorem providers a necessary condition on the transfer price to ensure that a Nash equilibrium exists with positive prices $\{p^*_i\}$, and both networks 1 and 2 having users and the Content Provider also having users. The second part of the theorem provides sufficient conditions for the existence of an ϵ-equilibrium. Further, it provides sufficient conditions for networks 1 and 2 and the Content Provider to all have users. Finally, it provides necessary and sufficient conditions to ensure that the prices $\{p^*_i\}$ constitute a Nash equilibrium in which networks 1 and 2 and the Content Provider all have users.

Theorem III.9. Existence of a Nash Equilibrium.

1) *If a Nash equilibrium exists with positive prices $\{p^*_i\}$, with networks 1 and 2 and the Content Provider each having users and making positive profit, then the transfer price satisfies*

$$-\frac{2+br}{3+r(3+4\widehat{C}_1)} < t < \frac{2+br}{3+r+2br}. \quad (26)$$

2) *When* (26) *holds and the expression for p^*_1 in* (20) *is positive, then the $\{p^*_i\}$ given by* (20), (21), (22) *constitute an ϵ-equilibrium, all the prices $\{p^*_i\}$ are positive, and networks 1 and 2 and the Content Provider all have users.*

3) *Necessary and sufficient conditions for the $\{p^*_i\}$ in* (20), (21), (22) *to constitute a Nash equilibrium in which networks 1 and 2 and the Content Provider all have users is that t falls within a specified subinterval of* (26).

Sketch Proof. The candidate solution $\{p^*_i\}$ is a local optimum for each i. The requirement that the p^* induce a feasible solution with each user making a positive profit yields (26) and $p^*_1 \geq 0$. The remaining feasibility conditions give the second equality for t. If the conditions in (26) are satisfied, then either the $\{p^*_i\}$ constitute a Nash equilibrium, or they are such that either network 1 or 2 could improve its profits by deviating, in which case p^*_i is an ϵ-equilibrium. This proves parts 1 and 2 of the theorem. The proof of part 3 involves finding, for each i, conditions on $b, \widehat{C}_1, r, t$ such that $\pi^*_i(t)$ is a Nash equilibrium.

The specific conditions, as well as characterization of ϵ, are given in the Supplementary Material. □

Corollary III.10. *When* (26) *holds, and $t \leq 0$ or $b > 1+\widehat{C}_1$, then the $\{p^*_i\}$ given by* (20), (21), (22) *constitute an ϵ-equilibrium.*

Sketch Proof. From (20) and the conditions of the corollary, we obtain $p^*_1 > 0$. The result follows from part 2 of the theorem. □

If the conditions of Theorem III.9 do not hold, there still remains the possibility of degenerate equilibria. We call a Nash equilibrium *degenerate* when at least one of the firms has no users, or at least one of the firms sets its price at zero. These—the only remaining possibilities for equilibria—are characterized by the following theorem.

Theorem III.11. Degenerate Equilibria. *There are only three possibilities for degenerate equilibria. Specifically, there exists a value t^A (which can be computed) such that:*

1) *If $t \geq t^A$, there exists a Nash equilibrium in which the Content Provider prices itself out of the market by setting $p^*_3 = t$, $Q^*_{13} = 0$. The network prices are*

$$p_1{}^* = \frac{2+r+r\widehat{C}_1}{3r\widehat{C}_1}, \quad p_2{}^* = \frac{1+2r+2r\widehat{C}_1}{3r\widehat{C}_1} \quad (27)$$

the user masses are

$$Q^*_1 = \frac{2+r+r\widehat{C}_1}{3(1+r+r\widehat{C}_1)}, \quad Q^*_2 = \frac{1+2r+2r\widehat{C}_1}{3(1+r+r\widehat{C}_1)} \quad (28)$$

and the profits are

$$\pi^*_1 = \frac{(2+r+r\widehat{C}_1)^2}{9r\widehat{C}_1(1+r+r\widehat{C}_1)}, \quad \pi^*_2 = \frac{[1+2(1+\widehat{C}_1)r]^2}{9r\widehat{C}_1(1+r+r\widehat{C}_1)}. \quad (29)$$

2) *If $t < 0$, network 1 provides a per-unit subsidy $s := -t$ to the Content Provider. If $s \geq \frac{2+br}{3+r(3+4\widehat{C}_1)}$, then there is a unique Nash equilibrium where the Content Provider sets its price $p^*_3 = 0$ and $Q^*_{11} = 0$. There are two subcases:*

   a) *If $s \leq \frac{2+br}{r\widehat{C}_1}$, then the network prices are*

$$p^*_1 = \frac{2+r(b+2\widehat{C}_1 s)}{3\widehat{C}_1 r}, \qquad p^*_2 = \frac{1+r(2b+\widehat{C}_1 s)}{3\widehat{C}_1 r} \quad (30)$$

   the user masses are

$$Q^*_{13} = \frac{2+r(b-\widehat{C}_1 s)}{3(1+br)}, \qquad Q^*_2 = \frac{1+r(2b+\widehat{C}_1 s)}{3(1+br)} \quad (31)$$

and the profits are

$$\begin{aligned}\pi_1 &= \frac{[2+r(b-\widehat{C}_1 s)]^2}{9r\widehat{C}_1(1+br)},\\ \pi_2 &= \frac{[1+r(2b+\widehat{C}_1 s)]^2}{9r\widehat{C}_1(1+br)}, \qquad (32)\\ \pi_3 &= \frac{[2+r(b-\widehat{C}_1)]s}{3(1+br)}.\end{aligned}$$

b) If $s > \frac{2+br}{r\widehat{C}_1}$, then network 1 chooses $p_1^* \geq s$. Consequently, $Q_2^* = 1$, and $p_1^* = s$, $p_2^* = s - \frac{1}{r\widehat{C}_1}$, with network 2 capturing all the profit, $\pi_2^* = s - \frac{1}{r\widehat{C}_1}$.

3) *If $-\frac{2+br}{3+r(3+4\widehat{C}_1)} < t < t^A$, there exists a set of parametric conditions, which includes $t > 0$, under which a Nash equilibrium exists where network 1 sets $p_1^* = 0$.*

Sketch Proof. No Nash equilibrium is possible in Region 2, since the feasibility condition will be $p_2 > p_1 + p_3 + p_3 + b(1-p_3)/\widehat{C}_1 > 0$, and network 2 can decrease its price until equality holds, attracting users and moving out of Region 2 into Region 3. If t is very large, the Content Provider will be shut out of the market, since $p_3 \geq t$ implies $Q_{13} = 0$, and we are in Case 1 of Lemma III.1. The equations for non-degenerate Case 1, and the first order conditions for the profits, give part 1 of the theorem. A sufficient condition for the solution of these first order conditions to constitute a Nash equilibrium is $t \geq t_A$ where t^A depends upon $b, \widehat{C}_1, r$ (and satisfies $t^A \leq 1$); t^A is given explicitly in the Supplementary Material.

Clearly there are potential Nash equilibria satisfying the conditions of Part 2 or 3 of the theorem; their existence under certain conditions is proved in the Supplementary Material.

Part 3 requires $\pi_1 \geq 0$. A necessary condition for existence here is that the transfer price is positive. Requiring the Nash prices to be optimal for networks 2 and the Content Provider requires that network 1 has sufficient capacity to carry all the content traffic, $\widehat{C}_1 > b + 1/r$; equivalently, the maximum amount users are prepared to pay for service exceeds the sum of the maximum possible congestion costs of networks 1 and network 2. The remaining condition to ensure a Nash equilibrium lead to constraints on t, detailed in the Supplementary Material. For fixed b, c, r satisfying $\widehat{C}_1 > b + 1/r$, the necessary and sufficient conditions for existence of this degenerate equilibrium reduce to t lying in a given positive interval.

Finally, there cannot be a Nash equilibrium under any of the remaining boundary conditions, all of which must be in Region 3. Since the profit functions $\pi_i(p_i)$ are strictly concave in Region 3, the only possible Nash equilibria in this case occur either at a unique interior point of the feasible region or at the boundaries of their support. It can be shown that no degenerate Nash can exist at these boundary points. □

Discussion. The theorem can be summarized as follows. There are three circumstances, which depend on the transfer price, for which a degenerate Nash equilibrium can exist. First, when the transfer price is so high (e.g., $t_A = 1$) that the Content Provider has no alternative but to charge a price that is sufficiently high as to price itself out of the market.

Second, if the transfer price is negative, i.e., is in fact a subsidy, and that subsidy is sufficiently great, then there will be a unique Nash equilibrium where the Content Provider sets its price to zero, and no users choose to take basic service from network 1. There are two subcases. If the subsidy, while sufficiently high, does not exceed the bound $\frac{2+br}{r\widehat{C}_1}$, then there will be users choosing network 1 for *additional* service via the Content Provider. If however the subsidy exceeds this bound, then there will be no users whatsoever choosing the content in this way (and hence all the users go to network 2).

When the transfer price t has a positive "intermediate" value then, as in Theorem III.9 (Existence of a Nash Equilibrium), it is possible that no Nash equilibrium exists, and only an ϵ-equilibrium exists, as seen from the following example.

Example III.12. *Consider the case $\widehat{C}_1 = 1$, $r = 3$, $b = 4$, $t = 1/4$, which violates the conditions necessary for p_2^* to be optimal. Using* (20), (21), (22) *gives the candidate Nash equilibrium point:*

$$\begin{aligned}\{p_i^*\} &= \frac{265}{162}, \frac{793}{324}, \frac{10}{27} &&\approx 1.64,\ 2.45,\ 0.37\\ \{\pi_i^*\} &= \frac{25477}{34992}, \frac{48373}{34992}, \frac{91}{11664} &&\approx 0.73,\ 1.38,\ 0.0078.\end{aligned}$$

These are optimal responses for network 1 and the Content Provider. However, network 2 can lower its price, and increase its return: if network 2 lowers its price to $589/324 \approx 1.82$, then it shuts the Content Provider out of the market, and increases its own profit to $346921/244944 \approx 1.42$. Hence the candidate solution is not a Nash equilibrium, only an $\epsilon-$equilibrium, where $\epsilon = \frac{1385}{40824} \approx 0.034$. With these values of $b, \widehat{C}_1, r, t$, there is a unique minimum-price best response by each network to the prices of the other two networks. However, this does not lead to a fixed point when the responses are iterated.

F. In the absence of a transfer price

There can be circumstances under which there is no transfer price (equivalently, $t \equiv 0$). In that case, the necessary and sufficient conditions for the existence of a Nash equilibrium remain complex. However, there exists a simple sufficient condition:

Theorem III.13. Existence of a Nash Equilibrium (no transfer price). *In the absence of a transfer price, a sufficient condition for a Nash equilibrium to exist is $b \leq 2(1+\widehat{C}_1) + 1/r$.*

Sketch Proof. This theorem, and subsequent Corollary, is a special case of Theorem III.9 (Existence of a Nash Equilibrium) where $t = 0$. When $t = 0$, the conditions for profits π^* and prices p^* to be non-negative are always satisfied, as is the condition for p_1^* to be optimal, leaving only the condition p_2^* to be optimal. However, a sufficient conditions for p_2^* to be optimal is $b \leq 2(1 + \widehat{C}_1) + 1/r$. □

In particular, the condition of the theorem is satisfied whenever the relative bandwidth is less than twice the effective capacity of network 1, or is less than 2, which would hold for many realistic settings.

The condition will fail to hold in situations where the bandwidth of the content is very high compared to "normal" service ($b \gg 1$), *and* the capacity of the enhanced network is much greater than that of the "normal" network ($r \gg 1$). In fact, there is a simple sufficient condition for a Nash equilibrium *not* to exist, as shown in the following corollary (for proof, see Supplementary Material).

Corollary III.14. *In the absence of a transfer price, a sufficient condition for a Nash equilibrium* not *to exist is* $b > 4(1+\widehat{C}_1) + 1/r$ *and* $r > 1$.

In particular, there will not exist a Nash equilibrium whenever the relative bandwidth is greater than four times the effective capacity of network 1, and the effective capacity of network 2 is greater than the effective capacity of network 1.

Example III.15. *Consider the case* $\widehat{C}_1 = 1$ *and* $C_2 = C_1$*, hence* $r = 1/b$*, when there is no transfer price. In this case the profits of the three firms are given by* $\frac{9b(2+3b)^2}{128(1+b)^2}, \frac{b(10+7b)^2}{128(1+b)^2}, \frac{9b(2+b)}{128(1+b)^2}$ *respectively. This example can be compared with what occurs when there is no Content Provider, when from* (29) *we have* $\pi_1^* = \frac{4(1+b)^2}{9(2+b)}$, $\pi_2^* = \frac{(4+b)^2}{9(2+b)}$. *The introduction of the Content Provider decreases the difference in profits between networks 1 and 2, increases profits for both networks 1 and 2 when* $b > 2$*, and decreases profits for both networks when* $b < 2$*. When* $b = 2$*, the introduction of a Content Provider leaves the profits of networks 1 and 2 unaltered (and identical, equal to 1), with the Content Provider making a small profit (of 1/16).*

In the case where there is no transfer price, we show that no firm can dominate the market:

Theorem III.16. User Masses (no transfer price). *In the absence of a transfer price, when a Nash equilibrium exists, at least one-quarter of users are carried on network 1 (including those choosing enhanced service from the Content Provider), but neither of the two types of users on network 1 can constitute a majority of all users. At least one-twelfth of users choose enhanced service from the Content Provider, and at least one-third choose enhanced service from network 2.*

In detail, we have the bounds:

$$
\begin{aligned}
0 <\ & Q_{11}^* < 1/2 \\
1/12 <\ & Q_{13}^* < 1/2 \\
1/3 <\ & Q_2^* < 3/4.
\end{aligned}
$$

Proof. Proof Putting $t = 0$ in (23) gives

$$
\begin{aligned}
Q_{11}^* &= \frac{br + 2}{2r(b + 2\widehat{C}_1 + 2) + 6} \\
Q_{13}^* &= \frac{(br + 2)(\widehat{C}_1 r + r + 1)}{2(br + 1)(r(b + 2\widehat{C}_1 + 2) + 3)} \\
Q_2^* &= \frac{br^2(b + 3\widehat{C}_1 + 3) + 2r(2b + \widehat{C}_1 + 1) + 2}{2(br + 1)(r(b + 2\widehat{C}_1 + 2) + 3)}.
\end{aligned}
$$

The bounds on the Q_i^* follow in a straightforward way, by considering limiting behavior under the constraints $b \geq 1$, $r > 0$ and $\widehat{C}_1 > 0$. (For example, the Q_2^* bounds follow by taking the limit as $r \downarrow 0$ or $r \uparrow \infty$). The lower bound for Q_{13}^* of 1/12 follows by noting that Q_{13} is decreasing in b, and putting $b = 4(1 + \widehat{C}_1) + 1/r$, since from Corollary III.14 a necessary condition for a Nash equilibrium to exist is $b \leq 4(1+\widehat{C}_1)+1/r$ or $r \leq 1$. □

IV. Optimizing the Transfer Price

We now allow network 1 to choose the transfer price. This reflects a regulatory setting where there are no restrictions on network 1 charging the Content Provider for access to its network.

The choice of transfer price is modeled as a Stackelberg leader-follower game [14]:

- First: Network 1 sets and publicly announces a transfer price, which may be positive or negative (i.e., a subsidy).
- Then: Network 2 and the Content Provider set their prices to maximize their respective revenues. Network 1 simultaneously sets its network price to maximize its revenue.

The solution to this game, called the *Stackelberg-Nash equilibrium*, is the vector $(t^*, p_1(t^*), p_2(t^*), p_3(t^*))$.

The following theorem shows that, when network 1 has the ability to choose the transfer price in this setting, there exists a unique Stackelberg-Nash equilibrium, where the form of the equilibrium depends upon the relationship among b, r and $\widehat{C}_1$.

Theorem IV.1. Uniqueness of the Transfer Price. *There exists a unique solution to the transfer price Stackelberg game. There are two cases: Either all three parties make a positive profit, or the two networks make positive profit and the Content Provider is shut out of the market. In both cases, network 1 sets a strictly positive transfer price* t^*.

Sketch Proof. We consider two subcases: $b \leq 1 + \widehat{C}_1$ and $b > 1 + \widehat{C}_1$. When $b \leq 1 + \widehat{C}_1$, recall that in this case the profit for network 1 is smaller with $t = 0$ than without the Content Provider. On calculating the profits from (7) using the prices and user masses from Theorem III.5 (Uniqueness of the Nash Equilibrium), we can see that $\pi_1(p_1^*(t), p_2^*(t), p_3^*(t))$ is quadratic in t. By direct calculation, it is straightforward to see that

$$
\frac{\partial^2}{\partial t^2}\pi_1(p_1^*(t), p_2^*(t), p_3^*(t)) < 0
$$

for $1 \leq b \leq 1 + \widehat{C}_1$. Hence any local optimum satisfying the first order conditions will be a global optimum within the feasible region of Case 3. Solving the first order conditions gives t^O as the solution to

$$
\frac{\partial}{\partial t}\pi_1(p_1^*(t), p_2^*(t), p_3^*(t)) = 0.
$$

Either this is the optimal value, or under certain conditions network 1 can be better off raising the transfer price to shut out the Content Provider, and hence will set a value $t \geq t^A$ and Theorem III.11 (Degenerate Equilibria), applies.

When $b > 1 + \widehat{C}_1$ and when a Nash equilibrium exists in Case 3, i.e., in Region 3, it is always advantageous for network 1 to have the Content Provider use its network. That is, the profit for network 1 is greater with $t = 0$ than with shutting the Content Provider out of the market. It is straightforward to show that, given $b > 1 + \widehat{C}_1$, it follows that $\frac{\partial}{\partial t}\pi_1(p_1^*(t), p_2^*(t), p_3^*(t))|_{t=0} > 0$. Hence, if a Nash equilibrium exists for a given value of t, then t should be strictly positive. In this case (i.e., Nash equilibrium and in Region 3), there are two possible cases

1) t can be $t^* = t^O$, i.e., solution to first order conditions
2) $t^* = t^B$ is on the boundary of the Nash equilibrium boundary, the critical point at which network 2 is indifferent between competing with the Content Provider, or lowering its price to drive out the Content Provider.

More precise specification of conditions under which the alternative holds are given in the following theorem, which is a compilation of several results proved earlier in the paper. □

Theorem IV.2. Specification of the Transfer Price. *The two cases for the transfer price, described in Theorem IV.1, can be specified as follows:*

1) *The prices, user masses, and profits are given by Theorem III.5 (Uniqueness of the Nash Equilibrium) with $t = t^B$or t^O, where:*
   a) *t^O solves the first-order profit maximization conditions This reduces to an affine equation in t as function of b, $\widehat{C}_1$ and r*
   b) *t^B is the point at which network 2 is indifferent between competing with the Content Provider or lowering its price to drive it out of the market.. (Here t^B naturally satisfies the feasibility constraint $t^B \leq \frac{2+br}{3+r+2br}$.)*
   Sufficient conditions for this case to exist are that $b^*(r, \widehat{C}_1) \leq b \leq 1+\widehat{C}_1$ OR $1+\widehat{C}_1 < b \leq 2(1+\widehat{C}_1)+1/r$ where b^* is the root of a cubic.
2) *The prices, user masses and profits are given by Theorem III.11 (Degenerate Equilibria), with $t^* \geq t^A$, where t^A is defined in Theorem III.11. There are two subcases:*
   a) *b is small, satisfying $1 \leq b \leq b^*(r, \widehat{C}_1)$ where b^* is the root of a cubic equation, with the bound $b^* < 1 + \widehat{C}_1$.*
   b) *b is large, for which sufficient conditions are $b > 4(1+\widehat{C}_1) + 1/r$ and $r > 1$ (c.f. Corollary III.14).*

Discussion. These two theorems show that is never optimal for network 1 to set a negative transfer price in the hope of attracting more revenue from the extra customers desiring enhanced service. There are instances where network 1 is better off using a subsidy (i.e., a negative transfer price) rather that setting a zero transfer price: but in each such instance, network 1 can do even better by raising the transfer price to such a level as to shut the Content Provider out of the market.

We have also shown that it is never optimal to set a zero transfer price, since $t^* > 0$.

The degenerate case, part 2, occurs when b is either small or is large. When b is small, less than $1 + \widehat{C}_1$, (subcase 2a), network 1 makes strictly greater profit than having a small positive transfer price that allows the Content Provider to compete. When b is large and r is greater than 1 (subcase 2b), network 1 cannot attain the higher profits that could otherwise be gained by not shutting out the Content Provider. Subcase 2b involves a social dilemma: network 1 (and network 2) would be better off at one of the Pareto optimal solutions where a small non-negative transfer price is set (e.g. $t = 0$) and all three parties are involved; however at any such transfer price, there is no Nash equilibrium in prices (e.g. at vector of prices $(p_i^*(0))$ network 2 has an incentive to lower its price to shut the Content Provider out).

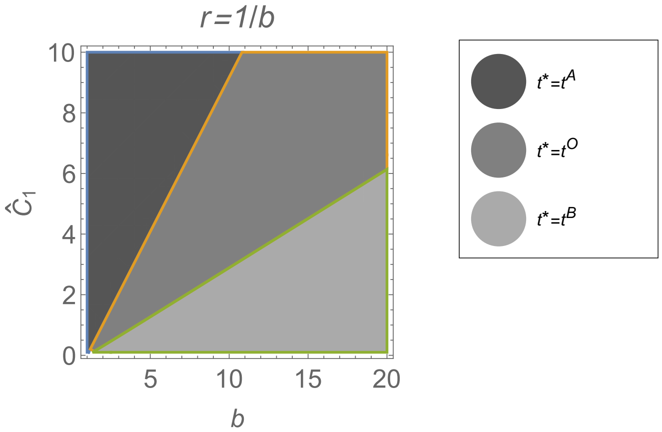

Fig. 2. Nash equilibria when $r = 1/b$ as $b, \widehat{C}_1$ vary.

A. Example: The case of equal capacities: $r = 1/b$

Consider the case when $r = 1/b$, corresponding to the two networks having equal capacity C. In this case the social dilemma instance of preemptive pricing does not exist.

The three regions corresponding to the different equilibria are shown in Figure 2, and the optimal choice of t is shown in Figure 3 when, in addition, $b = 10$. In the latter figure we show all possible values of $t^* \geq t^A$ that correspond to "shut-out values" (the shaded, rectangular region of the graph), rather than just $t^* = t^A$.

The transitions between the regions $t^* = t^O$ and $t^* = t^B$ is smooth (the value of t^* is uniquely defined on the boundary), whereas the transition between $t^* = t^O$ and the shut-out region (labeled $t^* = t^A$ in Figure 2) is not smooth: there is a phase transition at this point, as can be clearly seen in Figure 3. For this special case, direct calculation and algebraic manipulation shows that sufficient conditions for $t^* = t^O$ are

$$1 + \widehat{C}_1 \;\leq\; b \;\leq\; 1 + 3\widehat{C}_1 \tag{33}$$

and $t^O(b, \widehat{C}_1)$ increases with b and decreases with $\widehat{C}_1$ in this region.

V. Loosening the Assumptions

We now discuss what happens when we relax the assumptions of our base model. Specifically, we consider when (i) when the sensitivity to congestion is service dependent (ii) the congestion function is non-linear. We show that our assumption of making the congestion sensitivity (g) independent

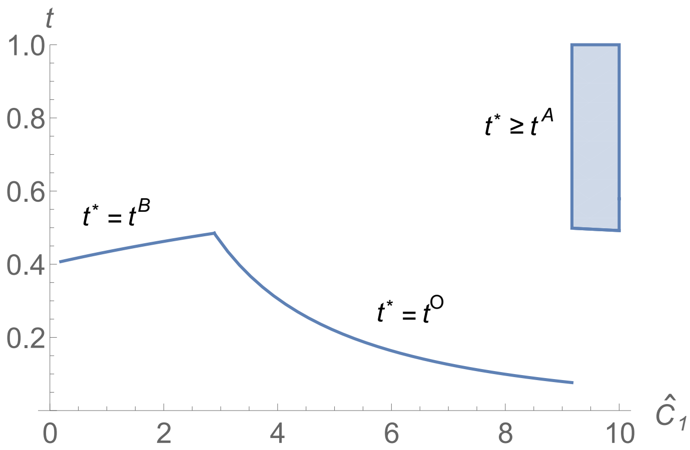

Fig. 3. Optimal t^* when $r = 1/b$, $b = 10$ as $\widehat{C}_1$ varies.

of the service is essentially without loss of generality. However, allowing non-linear congestion functions has nuanced implications; in particular, the optimal transfer price t^* and accompanying resulting *type* of equilibria depends upon the congestion function, producing qualitatively different behavior according to whether the convex congestion function is "close" to linear or steeper.

A. Service Dependent Sensitivity to Congestion

Allowing the sensitivity to congestion to be service dependent leads to qualitatively similar results. Suppose that g^B and g^E denote the sensitivity to congestion for the basic service and enhanced service, respectively. Then the user utilities would be modified accordingly. Define the "relative sensitivity" by $g = g^E/g^B$, and modify the effective capacities in the obvious way, i.e., $\widehat{C}_i = C_i/b^B g^B$, with the effective capacity ratio r defined in terms of these new quantities. Then Lemma III.1 still holds, with the critical value ∇_{12} defined as before, but now $\nabla_{13} = p_3 + (g-1)[Q_{11} + bQ_{13}/\widehat{C}_1]$. Hence for a given user mass on network 1 and a given price of content from the Content Provider, when the relative sensitivity exceeds one, more users prefer basic from service network 1 to going with the Content Provider than they would with a relative sensitivity of precisely one. The qualitative results of Section III that were derived previously about Nash equilibria continue to hold, but now there will be a dependence upon the relative sensitivity. For example, the prices in Theorem III.5 (Uniqueness of the Nash Equilibrium) will now depend on the relative sensitivity; however, the Nash equilibrium with positive prices will still be unique, and as before users masses will be linear functions of the prices.

B. Non-linear congestion costs

Consider the case of a general congestion function $k : \mathbb{R}_{\geq 0} \mapsto \mathbb{R}_{\geq 0}$, where k is convex and strictly increasing (instead of linear). Using the notation of Section III, writing K_i as shorthand for $K_i(\rho_i)$,

$$K_1 \;=\; K\left(\frac{Q_{11} + bQ_{13}}{\widehat{C}_1}\right), \quad K_2 \;=\; K\left(\frac{Q_2}{r\widehat{C}_1}\right) \qquad (34)$$

where $K(\rho) = k(g\rho)$. We can still write $\nabla_{12} = (K_2 - K_1) + p_2 - p_1$, $\nabla_{13} = p_3$ and Lemma III.1 holds under this mapping. Moreover, Theorem III.2 (Existence and Uniqueness of User Equilibrium) is also true.

When $t = 0$, as for the case of linear congestion functions, the only possible Nash equilibria are when all prices are positive, and each user makes a positive profit. When $t \neq 0$, then degenerate equilibria can exist; for example, as in Theorem III.11, when the transfer price is sufficiently large (part 1), a degenerate equilibrium exists when the Content Provider is shut out of the market. In this degenerate case, the expressions for the prices, and user masses become considerably more complicated, and depend on the function K, however qualitative behavior is the same (see below).

C. Non-linear congestion costs and no transfer price

For linear costs, uniqueness of the Nash equilibrium in Theorem III.5 followed from concavity of the profit functions. Concavity is a sufficient condition for uniqueness, and we now investigate how the shape of profit functions is affected by non-linear congestion costs.

It follows by differentiating the implicit equation defining Case 3, $(K_2 - K_1) + p_2 - p_1 = p_3$, that when $t = 0$ for any (increasing) function K, in Case 3,

$$0 > \frac{\partial Q_1}{\partial p_1} = \frac{\partial Q_2}{\partial p_2} = \frac{\partial Q_{13}}{\partial p_3} + \frac{K_2' + rK1'}{K_2' + brK1'}, \qquad (35)$$

where $K_1' := K'(\rho_1)$, $\rho_1 := \frac{Q_{11} + bQ_{13}}{\widehat{C}_1}$, with K_2' defined analogously; it also follows that when the first order conditions are satisfied, the user masses are related to prices via (24).

It also follows from the implicit equation that

$$\frac{\partial^2 \pi_1}{\partial p_1^2} = \frac{\partial Q_1}{\partial p_1}\left[2 + \left(b^2 K_1'' - \frac{K_2''}{r^2}\right)\frac{p_1}{\widehat{C}_1^2}\left(\frac{\partial Q_1}{\partial p_1}\right)^2\right] \qquad (36)$$

and hence π_1 will be a concave function of p_1 when $b^2 K_1'' - K_2''/r^2 > -\epsilon$ for a suitably defined small $\epsilon > 0$. At a turning point (p_i^*),

$$\frac{\partial^2 \pi_1}{\partial {p_1^*}^2} = -\frac{1}{p_1^* + p_2^*}\left[2 + \left(b^2 K_1'' - \frac{K_2''}{r^2}\right)\frac{1}{\widehat{C}_1^2}\frac{p_1^*}{(p_1^* + p_2^*)^2}\right] \qquad (37)$$

and hence we can derive sufficient conditions for this to be a local optimum by substituting for the allocations Q^* and calculating K_i''explicitly. Since

$$\frac{\partial^2 Q_1}{\partial p_1^2} = -\frac{\partial^2 Q_2}{\partial p_2^2} \qquad (38)$$

similar remarks apply to π_2 as a function of p_2.

For π_3, we can show

$$\frac{\partial^2 \pi_3}{\partial p_3^2} = \frac{p3}{\widehat{C}_1^2}\left(\frac{K_2''}{r^2} - K_1''\right) + \qquad (39)$$

$$\frac{\partial Q_{13}}{\partial p_3}\left[2\frac{p3}{\widehat{C}_1^2}\left(\frac{K_2''}{r^2} - bK_1''\right) + 2 + \right. \qquad (40)$$

$$\left. + \frac{p3}{\widehat{C}_1^2}\frac{\partial Q_{13}}{\partial p_3}\left(\frac{K_2''}{r^2} - b^2 K_1''\right)\right] \qquad (41)$$

and so sufficient conditions for π_3 to be concave function of p_3 are when $|b^2 K_1'' - K_2''/r^2| < \epsilon$ and $|K_1'' - K_2''/r^2| < \epsilon$ for small ϵ.

To summarize, unlike the linear case (when $K'' \equiv 0$) the profit functions are not necessarily concave, but will be around the point $b^2 K_1'' = K_2''/r^2$ (with an additional condition to ensure π_3 is semi-concave) , which for bounded congestion costs implies when $r \approx 1/b$. Hence in these circumstances the (non-degenerate) Nash equilibrium of Theorem III.5 will be unique.

D. Quadratic congestion costs

Let us look at the quadratic congestion case, $K(\rho) = \kappa\rho^2$, in more detail. We can put $\kappa = 1$ without loss of generality, by redefining $\widehat{C}_1 = \frac{\sqrt{\kappa}C_1}{b^B g}$. When $r = 1/b$ and we are in Case 3 (all networks make a positive profit), we have shown that π_1 and π_2 are concave functions (of p_1, p_2 respectively). Refining the analysis given above, we can show that π_3 is a concave function when $r = 1/b$ provided $b < \sqrt{3 + \widehat{C}_1^2}$, and will be quasi-concave if $p_3 < 2/3$.

Now Theorem III.5 needs to be amended to say that there are sufficient conditions for a unique Nash equilibrium (i.e., it will be if $r = 1/b$, $b < \sqrt{3 + \widehat{C}_1^2}$ or $p_3^* < 2/3$), but the corresponding expressions for p^* are complicated.

Corresponding to Theorem III.11, there will be a degenerate equilibrium if $t \geq t^A$ shutting out the Content Provider. Here, the qualitative behavior is the same as under linear congestion costs. For example, taking $K(\rho) = \rho^2$ and putting $r = 1/b$, we can show that network 1 has at least 1/3 of the users and at most 60% of users (hence network 2 has at least 40% of users), whereas for linear congestion, taking the limits in Theorem III.11, part 1, network 1 takes between 1/3 and 2/3 of the users.

There also two degenerate cases corresponding to parts (2) and (3) of Theorem III.11.

However, when optimizing over t, we have *different* behavior to the linear congestion case. In particular, it can be optimal for network 1 to offer a subsidy to the Content Provider to increase its profits. As an example when $r = 1/b$, the case of equal capacity C, when $b = 10$ we show the optimal choice of t in Figure 4 which can be directly compared with Figure 3. The corresponding profits are shown in Figure 5.

What is happening is that for quadratic congestion, if network 1 sets a small negative transfer price rather than 0, this allows the Content Provider to lower its price, thereby attracting more traffic (increasing Q_{13}). The effect of this extra traffic will cause more congestion (if $p_3 + Q_{13} = Q_1$ increases) on network 1, which has the effect of raising equilibrium prices on networks 1 and 2; the increase in the product $p_1 Q_1$ is larger than the decrease tQ_{13}, so network 1 makes a higher profit. If the overall capacity of the networks $C_1 + C_2$ is comparable to the maximum enhanced service demand, gb^E, then it becomes more advantageous for network 1 to shut out the Content Provider, by setting a high transfer price.

With equal capacities and quadratic congestion, network 1's optimal strategy is simply described. When the capacity is small compared to b, its optimal strategy is to set a negative

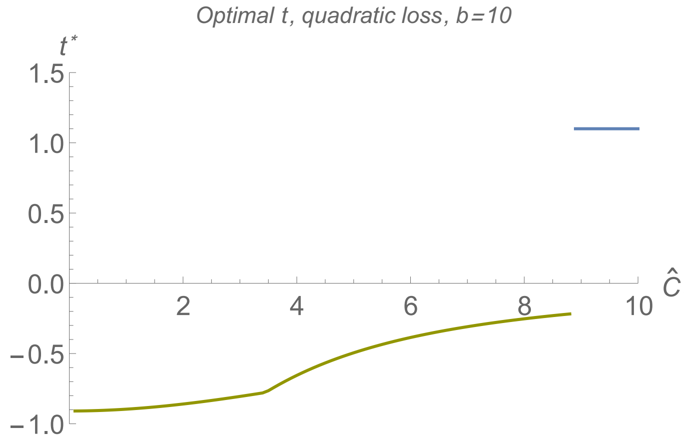

Fig. 4. Optimal t^* a when $r = 1/b$, $b = 10$ as $\widehat{C}_1$ varies.

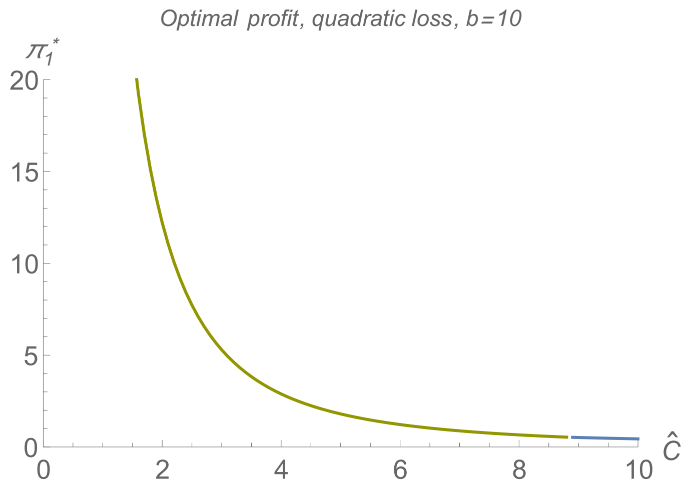

Fig. 5. Optimal π_1^* a when $r = 1/b$, $b = 10$ as $\widehat{C}_1$ varies.

transfer price, i.e., subsidize the Content Provider; when the capacity is large compared to b, network 1's optimal strategy is to shut out the Content Provider by setting a high transfer price and sharing the profits with network 2. More formally, we have:

Theorem V.1. Quadratic Costs and Equal Capacities. *For quadratic congestion costs and equal capacities, if network 1 does not shut out the Content Provider with a high price, then network 1's optimal strategy is to set a negative transfer price* $t^* = t^O(b, \widehat{C}_1)$.

Sketch Proof. The proof consists of showing that $\frac{d}{dt}\pi_1^*(t)|_{t=0} < 0$, and that $\pi_1^*(t)$ is concave, or that $\frac{d}{dt}\pi_1^*(t) < 0$ for all $t \geq 0$. This will establish that setting a shut out value $t^* = t^A$ becomes increasingly attractive as $\widehat{C}_1$ increases; that is, the revenue benefit of setting a shut out value over setting a transfer price increases with $\widehat{C}_1$. When in Case 3 of Lemma III.1, putting $r = 1/b$, using $Q_{11} = p_3$, and substituting in the defining conditions, gives Q_{13} as the solution to a quadratic equation, which reduces to a linear equation with solution

$$Q_{13} = \frac{b^2(1-p_3)^2 - p_3^2 - \widehat{C}_1^2(p_1 - p_2 + p_3)}{2b(b(1-p_3)+p_3)}.$$

We have seen that $\pi_1(p_1)$ and $\pi_2(p_2)$ are concave when $r = 1/b$, and that $\pi_3(p_3)$ is in general quasi-concave, hence we can find the optimum by considering the first order conditions. This leads to a set of equations that reduce to a cubic in p_3, where p_3^* is the real root in $[0, 1]$. Using Mathematica [18], one can show that $\frac{d}{dt}\pi_1^*(t)|_{t=0} < 0$. □

Technical Remark V.2. *We can also prove an analogous result to Theorem V.1 for the case $r = 1$ (rather than $r = 1/b$), i.e., again for quadratic congestion the optimal strategy is to subsidize the Content Provider. Recall that $r = 1$ corresponds to network 2 being able to provide all users with enhanced traffic at a level of congestion equal to that of network 1, were network 1 to carry all the users with basic service.*

E. General convex congestion costs

Consider a general superlinear congestion function $K(\rho) = \rho^{1+\alpha}$ for $\alpha \geq 0$. Thus, $\alpha = 0$ corresponds to linear congestion costs, and $\alpha = 1$ to quadratic. The general behavior for fixed t is similar to that for the linear or quadratic case; namely, there are conditions under which a non-degenerate Nash equilibrium exists, which is "in general" unique. Non-degenerate Nash equilibria exist, for example when the transfer price is sufficiently high as to shut out the Content Provider.

What happens when network 1 chooses t optimally? There are then two distinct scenarios. For given parameters b, r, and $\widehat{C}_1$, there is some value of $\alpha \in [1, 2]$, $\alpha^*(b, r, \widehat{C}_1)$, such that: (1) for $\alpha < \alpha^*$, the optimal decision shares the same property with linear congestion costs ($\alpha \equiv 0$), viz., either set a positive transfer price, or set the price so high so as to shut out the Content Provider; and (2) for $\alpha \geq \alpha^*$, the optimal decision shares the same property with quadratic congestion costs ($\alpha \equiv 1$), viz., either set a negative transfer price and thus subsidize the Content Provider, or set the price so high so as to shut out the Content Provider.

VI. Concluding Remarks

This paper considers an ISP providing basic Internet service (at price p_1) competing with an ISP that provides *enhanced service*, i.e., both Internet service and content (at price p_2), where the basic service ISP can partner with a Content Provider, who charges each user an additional price (p_3) for the content, where the Content Provider pays the basic service ISP a transfer price (t) for delivering the content. The transfer price can be positive, negative, or zero. A positive transfer price could be a termination fee reflecting discriminatory pricing by network 1 against the Content Provider; note that this would contravene the *zero-price rule* interpretation of net neutrality [19], [9]. Alternatively, the positive transfer price may be compensation paid to network 1 mandated by a regulator, or an agreed transfer price negotiated bilaterally between network 1 and the Content Provider.

This gives rise to questions of when a Nash equilibrium in prices will exist, whether the Nash equilibrium will be unique, and what prices will be charged by the networks and the Content Provider, and whether each of them will make a profit. When network 1 has the ability to choose the transfer price, there is the question of whether and when the transfer price should be negative (i.e., a subsidy),

We find that answers to these questions can be expressed concisely by defining three intermediate concepts. The first is the *effective capacity* of a network, $\widehat{C}_i$ $(i = 1, 2)$, which is the capacity of the network divided by product of the expected bandwidth per user and the user sensitivity to congestion on the network. The effective capacity can be interpreted as the mass of users that the network can tolerate before "saturating," taking into account not only bandwidth but also the service-dependent sensitivity to price. The second concept is that of the *effective capacity ratio*, r, which is the effective capacity of the ISP providing its own content (network 2) divided by the effective capacity the ISP providing only basic Internet service (network 1). The third is that of the *relative bandwidth*, b, the ratio of enhanced service bandwidth to basic service bandwidth.

Our results are as follows. When the transfer price is fixed and known *ex ante* by all three firms, then in a well provisioned network where the congestion is linear, there are closed-form necessary and sufficient conditions under which a Nash equilibrium will exist. If an equilibrium exists with positive prices where the two networks and the Content Provider all have users and make a profit, then the equilibrium will be unique. In this case, we can provide closed-form expressions for all three equilibrium prices, p_1, p_2, and p_3. We also provide bounds on the transfer price, t.

The relative bandwidth, b, plays a key role in the characterization of the equilibrium. This relationship can be seen most clearly when there is no transfer price, i.e., $t = 0$, in which case a sufficient condition for a Nash equilibrium to exist is $b \leq 2(1 + \widehat{C}_1) + 1/r$, where $\widehat{C}_i$ is the effective capacity of network i, and r is effective capacity ratio $\widehat{C}_2/\widehat{C}_1$. This condition holds whenever the relative bandwidth is less than twice the effective capacity of network 1, or is less than 2, which would be the case for many realistic settings. For example, when content is relatively low bit-rate stream, we would have $b \leq 2$.

The paper of Johari, Weintraub and Roy [8] is complementary to ours, in that it considers the question of investment as well as price setting, and for homogeneous rather than heterogeneous users. We can deal with homogeneous users in our model by setting the continuous preference space w to a point mass, w_0, and carrying through our analysis for this degenerate case. However, the presence of the Content Provider, which has no capacity of its own, means that in general user masses will no longer be uniquely defined. Nevertheless, when there is no transfer price and where $b = 1$, our results carry over with the appropriate adjustments, and we can show that, analogous to Theorem III.5, there exists a unique Nash equilibrium with networks 1 and 2 and the Content Provider each having users, provided that a simple condition is satisfied ($w_0\widehat{C}_1 \leq 1 + 2/r$). In this setting, the Content Provider has no choice other than to price at $p_3 = w_0$. For our homogeneous model we can characterize and establish the uniqueness of a pure-strategy Nash equilibrium, and we find that cost structure has a critical impact on market outcomes, both of which were results of the Johari, Weintraub and Roy [8] model.

When network 1 has the ability to choose the transfer price, we model the problem as a Stackelberg leader-follower game in which, first network 1 sets and publicly announces the transfer price, and then network 2 and the Content Provider set their prices to maximize their respective revenues, with network 1 simultaneously setting its network price. Here, the results are dependent upon the shape of the congestion function. If the congestion function is linear, as would likely be the case where both networks are well-provisioned, then it is never optimal for network 1 to set a negative transfer price in the hope of attracting more revenue from the additional customers desiring enhanced service. Further, it is optimal for network 1 to set a strictly positive transfer price; in other words, net neutral pricing is not optimal for network 1. The optimal transfer price uniquely determines one of two outcomes: either it enables all three parties to make a profit, or it shuts out the Content Provider, creating a duopoly. We provide necessary and sufficient conditions for each outcome to occur, and characterize the transfer price for each outcome.

When the congestion function is non-linear, the equilibria and qualitative results differ according to whether the convex congestion function is "close" to linear or whether it is steeper.

When the congestion function is quadratic, and the capacities of the two networks are the same, then if network 1 does not shut out the Content Provider with a high price, then the optimal strategy for network 1 is to subsidize the Content Provider, i.e., set a negative transfer price.

If, on the other hand, the congestion function is superlinear—up to quadratic—reflecting a possible restricted capacity in the network, then the general behavior for fixed transfer price t is similar to that of the linear or quadratic cases, depending on whether the congestion function is, respectively, closer to linear, or close to quadratic. However, the characterization of the optimal transfer price strategy changes. When the capacities of the two networks are the same, then the optimal strategy for network 1 is to either subsidize the Content Provider (set a negative transfer price) or shut out the Content Provider. We conjecture that this result for superlinear congestion costs extends to general convex congestion costs.

Finally, we have not considered investment incentives and market structure. This seems like an important area for future work, although the models are likely to be complex. A good place to begin would be the paper of Johari, Weintraub and Roy.

References

[1] S. Ramachandran and D. FitzGerald, "For web firms, faster access comes at a price," *Wall Street Journal (Online)* June 20., p. n/a, June 2013.

[2] N. Economides and B. E. Hermalin, "The economics of network neutrality," *The RAND Journal of Economics*, vol. 43, no. 4, pp. 602–629, Dec 2012.

[3] K. Finley, "Comcast's Netflix deal could open a new front in net neutrality war," *Wired*, Jul. 2016.

[4] A. de Palma and L. Leruth, "Congestion and game in capacity: a duopoly analysis in the presence of network externalities," *Annales d'Economie et de Statistique*, no. 15/16, pp. 389–407, 1989.

[5] J. R. Marden and A. Wierman, "Distributed welfare games," *Operations Research*, vol. 61, no. 1, pp. 155–168, Feb 2013.

[6] R. Gibbens, R. Mason, and R. Steinberg, "Internet service classes under competition," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 12, pp. 2490–2498, Dec 2000.

[7] N. Shetty, G. Schwartz, and J. Walrand, "Internet QoS and regulations," *IEEE/ACM Transactions on Networking*, vol. 18, no. 6, pp. 1725–1737, Dec 2010.

[8] R. Johari, G. Y. Weintraub, and B. V. Roy, "Investment and market structure in industries with congestion," *Operations Research*, vol. 58, no. 5, pp. 1303–1317, Oct 2010.

[9] R. T. B. Ma and V. Misra, "The public option: A nonregulatory alternative to network neutrality," *IEEE/ACM Transactions on Networking*, vol. 21, no. 6, pp. 1866–1879, Dec. 2013.

[10] R. T. B. Ma, "Subsidization competition: Vitalizing the neutral internet," *IEEE/ACM Transactions on Networking*, vol. 24, no. 4, pp. 2563–2576, Aug. 2016.

[11] A. Shaked and J. Sutton, "Relaxing price competition through product differentiation," *Review of Economic Studies*, vol. 49, no. 1, pp. 3–13, Jan 1982.

[12] ——, "Natural oligopolies," *Econometrica*, vol. 51, no. 5, pp. 1469–1483, Sep 1983.

[13] G. P. Cachon and P. Feldman, "Pricing services subject to congestion: Charge per-use fees or sell subscriptions?" *Manufacturing and Service Operations Management*, vol. 13, no. 2, pp. 244–260, Apr 2011.

[14] E. Altman, T. Boulogne, R. El-Azouzi, T. Jiménez, and L. Wynter, "A survey on networking games in telecommunications," *Computers & Operations Research*, vol. 33, pp. 286–311, Feb 2006.

[15] A. Rubinstein, *Modeling Bounded Rationality*. Cambridge, MA: MIT Press, 1998.

[16] R. Radner, "Collusive behavior in non-cooperative epsilon equilibria of oligopolies with long but finite lives," *Journal of Economic Theory*, vol. 22, no. 2, pp. 136–154, Apr 1980.

[17] V. Martínez-de-Albéniz and D. Simchi-Levi, "Competition in the supply option market," *Operations Research*, vol. 57, no. 5, pp. 1082–1097, Oct 2009.

[18] Wolfram Research, Inc., *Mathematica, version 10.4*, Champaign, Illinois.

[19] F. Schuett, "Network neutrality: A survey of the economic literature," *Review of Network Economics*, vol. 9, no. 2, Article 1, Jan 2010.

Supplementary Material
Proofs of Theorems and Corollaries

A. Proof of Lemma III.1

Using (11) with the definitions of ∇_{12} and ∇_{13} given in (12) and (13), a comparison of utilities from (2) and (3), (2) and (4), and (3) and (4), respectively, yields the following.

In equilibrium, a user with willingness to pay w will weakly prefer:

i) Network 1 to network 2 if $w \le \nabla_{12}$, network 2 over network 1 if $w \ge \nabla_{12}$

ii) Network 1 to option 3 if $w \le \nabla_{13}$, option 3 over network 1 if $w \ge \nabla_{13}$

iii) Network 2 to option 3 if $\nabla_{12} \le \nabla_{13}$, option 3 over network 2 if $\nabla_{12} \ge \nabla_{13}$, regardless of w in both cases, and strict preference obtains when the corresponding inequality is strict.

Consider Case 1 of Lemma III.1. By iii, all users prefer network 2 to option 3. Thus, any users taking option 3 would migrate from network 1 to network 2. Hence Q_{13} would decrease and Q_2 would increase, until either there is equality or until Q_{13} is zero. (The preference for network 1 or network 2 is given by i.)

In Case 2 of the Lemma, by iii all users prefer option 3 to network 2; thus, any users on network 2 would migrate to network 1 to choose option 3. Hence Q_2 decreases and Q_1 increases until either there is equality, or until Q_2 is zero.

For the equality of Case 3 of the Lemma, it follows from iii that at equilibrium all users are indifferent between network 2 and option 3, but clearly they will not all choose one or another, since otherwise some could reduce their cost by choosing the empty network.

B. Proofs of Theorem III.2 and Corollary III.3

1) Proof of Theorem III.2:

Proof. We prove uniqueness and existence of $\mathbf{Q}$ by explicitly characterizing Q_{11}, Q_{13}, Q_2 in Theorem III.2 in the three mutually exclusive cases of Lemma III.1. In each case, we first show that Q_{11}, Q_{13}, Q_2 are uniquely determined—specifically, the solution of an appropriate system of linear equations; we then establish feasibility, and hence the existence of the unique solution.

Case 1: No users join the Content Provider, because the price is too high, hence $Q_{13}=0$. From Lemma III.1, it follows that given a vector of prices $\mathbf{p}$, the vector $\mathbf{Q}$ must satisfy

$$Q_{13} = 0 \tag{S.1}$$

$$\nabla_{12} = \frac{1}{r\widehat{C}_1}(Q_2 - rQ_{11}) + p_2 - p_1 \tag{S.2}$$

$$Q_{11} = [\nabla_{12}]_0^1 \tag{S.3}$$

$$Q_2 = 1 - Q_{11}. \tag{S.4}$$

where $[x]_l^u$ denotes the function equal to x when $l \le x \le u$, equal to l when $x < l$, and equal to u when $x > u$. Solving (S.1)–(S.4) yields:

$$(Q_1, Q_2) = \begin{cases} (1,0) \\ \left(\frac{1+\widehat{C}_1(p_2-p_1)r}{1+r+\widehat{C}_1 r}, \frac{[1-\widehat{C}_1(p_2-p_1-1)]r}{1+r+\widehat{C}_1 r}\right) \\ (0,1) \end{cases} \tag{S.5}$$

if, respectively, $p_2-p_1 > 1+\frac{1}{\widehat{C}_1}$, $1+\frac{1}{\widehat{C}_1} \ge p_2-p_1 \ge -\frac{1}{r\widehat{C}_1}$, and $-\frac{1}{r\widehat{C}_1} > p_2-p_1$. The profits are:

$$\pi_1 = p_1 Q_1 \tag{S.6}$$

$$\pi_2 = p_2 Q_2 \tag{S.7}$$

$$\pi_3 = 0. \tag{S.8}$$

We have from (S.2) and (S.5):

$$\nabla_{12} = \begin{cases} -\frac{1}{\widehat{C}_1} + p_2 - p_1 & \text{if} \quad p_2-p_1 > 1+\frac{1}{\widehat{C}_1} \\ \frac{1+\widehat{C}_1(p_2-p_1)r}{1+r+\widehat{C}_1 r} & \text{if} \quad 1+\frac{1}{\widehat{C}_1} \ge p_2-p_1 \ge -\frac{1}{r\widehat{C}_1} \\ \frac{1}{r\widehat{C}_1} + p_2 - p_1 & \text{if} \quad -\frac{1}{r\widehat{C}_1} > p_2-p_1 \end{cases} \tag{S.9}$$

which correspond to the three regions of ∇_{12}, respectively, being less than 1, in the closed interval $[0,1]$, and smaller than 1.

Feasibility: From (14) and (15) and substituting into (S.6) and (S.7) gives the necessary and sufficient conditions: $p_1 \ge 0$ and $p_2 \ge 0$. From (S.9) and (13), and using the defining expression for Case 1, which is given in terms of the user masses and the prices, we have an alternative defining expression for Case 1 in terms of the prices alone:

$$p_3 > \begin{cases} -\frac{1}{\widehat{C}_1} + p_2 - p_1 & \text{if} \quad p_2-p_1 > 1+\frac{1}{\widehat{C}_1} \\ \frac{1+\widehat{C}_1(p_2-p_1)r}{1+r+\widehat{C}_1 r} & \text{if} \quad 1+\frac{1}{\widehat{C}_1} \ge p_2-p_1 \ge -\frac{1}{r\widehat{C}_1} \\ \frac{1}{r\widehat{C}_1} + p_2 - p_1 & \text{if} \quad -\frac{1}{r\widehat{C}_1} > p_2-p_1 \end{cases}. \tag{S.10}$$

Case 2: No users join network 2 because the price is too high. Using (13):

$$Q_{11} = [p_3]^1$$
$$Q_{13} = 1 - [p_3]^1$$

where $[x]^u$ denotes the function equal to x when $x \le u$, and equal to u when $x > u$. The requirement that firms 1 and 3 make non-negative profits, and the expressions for profits, (7), yields:

$$p_1 + t(1-[p_3]^1) \ge 0 \tag{S.11}$$

$$(p_3 - t)(1-[p_3]^1) \ge 0. \tag{S.12}$$

Feasibility: This region is feasible if the demand allocations are feasible and profits are non-negative, and in addition the defining condition for Case 2 holds. Using (S.12), (S.11), and the defining condition for Case 2, the conditions are:

$$p_3 \ge t \tag{S.13}$$

$$p_1 + t(1-[p_3]^1) \ge 0 \tag{S.14}$$

$$p_2 - p_1 > \frac{b + \widehat{C}_1 p_3 + (1-b)[p_3]^1}{\widehat{C}_1}. \tag{S.15}$$

Case 3: From (13) $\nabla_{13} = p_3$, and since $Q_{11} = [\nabla_{13}]^1$, we have:

$$Q_{11} = [p_3]^1 \tag{S.16}$$

and also have:

$$Q_2 \;=\; 1-(Q_{11}{+}Q_{13}). \tag{S.17}$$

Subcase A: $\mathbf{0 \le p_3 \le 1}$*:* Solving simultaneously the defining condition for Case 3, (S.16), (S.17) together with (6) and (10) yields:

$$Q_{11} = p_3 \tag{S.18}$$

$$Q_{13} \;=\; \frac{1+r\widehat{C}_1(p_2{-}p_1)-(1{+}r{+}r\widehat{C}_1)p_3}{1+br} \tag{S.19}$$

$$Q_2 \;=\; \frac{r[b-\widehat{C}_1(p_2{-}p_1)+(1{-}b{+}\widehat{C}_1)p_3]}{1+br} \tag{S.20}$$

$$Q_1 \;=\; \frac{1+r\widehat{C}_1(p_2{-}p_1)-r(1{-}b{+}\widehat{C}_1)p_3}{1+br}. \tag{S.21}$$

Feasibility: For the demand allocations to be feasible given the prices, we require (18) to hold. The nonnegativity conditions on the Q's given by (14), (15), and (16) imply necessary and sufficient conditions for Case 3 to be feasible. These conditions are that the p_i satisfy:

$$\begin{aligned}
0 \;&\le\; p_3 \;\le 1 && \text{(S.22)}\\
p_3 \;&\le \frac{1+r(p_2{-}p_1)\widehat{C}_1}{1+r+r\widehat{C}_1} && \text{(S.23)}\\
(p_2{-}p_1)\widehat{C}_1-b \;&\le\; (1{-}b{+}\widehat{C}_1)p_3 && \text{(S.24)}\\
0 \;&\le\; p_1\left[1+r(p_2{-}p_1)\widehat{C}_1-r(1{-}b{+}\widehat{C}_1)p_3\right]\\
&\quad +t\left[1+r(p_2{-}p_1)\widehat{C}_1-(1{+}r{+}r\widehat{C}_1)p_3\right] && \text{(S.25)}\\
0 \;&\le\; p_2\left[rb-r\widehat{C}_1(p_2{-}p_1)+r(1{-}b{+}\widehat{C}_1)p_3\right] && \text{(S.26)}\\
0 \;&\le\; (p_3-t)\left[1+r\widehat{C}_1(p_2{-}p_1)-(1{+}r{+}\,r\widehat{C}_1)p_3\right]. && \text{(S.27)}
\end{aligned}$$

Note that the defining condition for Case 3 is satisfied by the Q's by construction.

Note also that in the "fully non-boundary" case, when all the user masses are strictly positive, the conditions simplify to

$$\begin{aligned}
0 \;&<\; p_3 \;<1\\
p_3 \;&<\; \frac{1+r(p_2{-}p_1)\widehat{C}_1}{1+r+r\widehat{C}_1}\\
r[(p_2{-}p_1)\widehat{C}_1-b] \;&<\; r(1{-}b{+}\widehat{C}_1)p_3\\
-p_1\frac{1+r(p_2{-}p_1)\widehat{C}_1-r(1{-}b{+}\widehat{C}_1)p_3}{1+r(p_2{-}p_1)\widehat{C}_1-(1{+}r{+}r\widehat{C}_1)p_3} \;&\le\; t\\
0 \;&\le\; p_2\\
t \;&\le\; p_3\,.
\end{aligned}$$

Subcase B: $\mathbf{p_3 > 1}$*:* Solving simultaneously the defining condition for Case 3, (S.16), (S.17) together with (6) and (10) yields:

$$\begin{aligned}
Q_{11} \;&=\; 1 && \text{(S.28)}\\
Q_{13} \;&=\; 0 && \text{(S.29)}\\
Q_2 \;&=\; 0 && \text{(S.30)}\\
Q_1 \;&=\; 1. && \text{(S.31)}
\end{aligned}$$

Feasibility: These conditions are that the p_i satisfy:

$$\begin{aligned}
p_2{-}p_1{-}p_3 \;&=\; \frac{1}{\widehat{C}_1} && \text{(S.32)}\\
p_1 \;&\ge\; 0 && \text{(S.33)}\\
p_3 \;&>\; 1. && \text{(S.34)}
\end{aligned}$$

□

2) Proof of Corollary III.3:

Proof. The characterization of the constraints corresponding to (14) to (18), for each of the three cases of Lemma III.1, is given above in part 1, "Proof of Theorem III.2." When $t = 0$, each of the constraints corresponds to a separating hyperplane or the space formed by intersecting hyperplanes (e.g., (S.26) is equivalent to $p_2 \ge 0$ and (S.24)). The same holds true when $t > 0$, apart from the constraint for $\pi_1 \ge 0$ for Case 3, (S.25), which is quadratic in p_1 but which reduces to an intersection of hyperplanes. When $t < 0$, it is straightforward to show the region is convex. Combining these statements proves the corollary. □

C. *Proof of Theorem III.5*

The system {(S.18), (S.20), (S.19)} in matrix form is: $\widehat{C}_1(1{+}br)\mathbf{Q} = \mathbf{c} + \mathbf{q}{\cdot}\mathbf{p}$. Equivalently:

$$\widehat{C}_1(1{+}br)\begin{pmatrix} Q_{11}\\ Q_2\\ Q_{13}\end{pmatrix} = \begin{pmatrix} c_1\\ c_2\\ c_3\end{pmatrix} + (q_{ij}).\begin{pmatrix} p_1\\ p_2\\ p_3\end{pmatrix} \tag{S.35}$$

where

$$\mathbf{c} = \begin{pmatrix} 0\\ b\widehat{C}_2\\ \widehat{C}_1\end{pmatrix},\;\mathbf{q} = \begin{pmatrix} 0 & 0 & (1{+}br)\widehat{C}_1\\ \widehat{C}_1\widehat{C}_2 & -\widehat{C}_1\widehat{C}_2 & (\widehat{C}_1{-}b{+}1)\widehat{C}_2\\ -\widehat{C}_1\widehat{C}_2 & \widehat{C}_1\widehat{C}_2 & -(1{+}r{+}r\widehat{C}_1)\widehat{C}_1\end{pmatrix}. \tag{S.36}$$

Using (7) together with (S.35) gives:

$$\begin{aligned}
\widehat{C}_1(1{+}br)\begin{pmatrix}\pi_1\\ \pi_2\\ \pi_3\end{pmatrix} \;&=\; \widehat{C}_1(1{+}br)\mathbf{P}{\cdot}\mathbf{Q}\\
&= \begin{pmatrix} p_1 & 0 & p_1+t\\ 0 & p_2 & 0\\ 0 & 0 & p_3-t\end{pmatrix}\cdot(\mathbf{c}+\mathbf{q}{\cdot}\mathbf{p}).
\end{aligned}$$

Taking derivatives

$$\begin{aligned}
\widehat{C}_1(1{+}br)\begin{pmatrix}\frac{\partial\pi_1}{\partial p_1}\\ \frac{\partial\pi_2}{\partial p_2}\\ \frac{\partial\pi_3}{\partial p_3}\end{pmatrix} \;&=\; \begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\cdot(\mathbf{c}+\mathbf{q}\cdot\mathbf{p})\\
&+\begin{pmatrix} q_{11}+q_{31} & 0 & 0\\ 0 & q_{22} & 0\\ 0 & 0 & q_{33}\end{pmatrix}\cdot\mathbf{p}+\mathbf{t}\begin{pmatrix} q_{31}\\ 0\\ -q_{33}\end{pmatrix}.
\end{aligned} \tag{S.37}$$

Hence at the potential N.E. where $\frac{\partial \pi_i}{\partial p_i} = 0$ for all i, a simultaneous turning point, the p_i^* will satisfy

$$t\begin{pmatrix}-q_{31}\\0\\q_{33}\end{pmatrix} - \begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}\cdot\mathbf{c}$$
$$= \begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}\cdot\mathbf{q}\cdot\mathbf{p} + \begin{pmatrix}q_{11}+q_{31}&0&0\\0&q_{22}&0\\0&0&q_{33}\end{pmatrix}\cdot\mathbf{p}$$
$$= \begin{pmatrix}2(q_{11}+q_{31})&q_{12}+q_{32}&q_{13}+q_{33}\\q_{21}&2q_{22}&q_{23}\\q_{31}&q_{32}&2q_{33}\end{pmatrix}\cdot\mathbf{p}$$
$$= \tilde{\mathbf{q}}\cdot\mathbf{p} \tag{S.38}$$

Now it follows from (S.37) that

$$\begin{pmatrix}\frac{\partial^2\pi_1}{\partial p_1^2}\\\frac{\partial^2\pi_2}{\partial p_2^2}\\\frac{\partial^2\pi_3}{\partial p_3^2}\end{pmatrix} = 2\begin{pmatrix}q_{11}+q_{31}&0&0\\0&q_{22}&0\\0&0&q_{33}\end{pmatrix} \tag{S.39}$$
$$= 2\begin{pmatrix}-\widehat{C}_1\widehat{C}_2&0&0\\0&-\widehat{C}_1\widehat{C}_2&0\\0&0&-(1+r+r\widehat{C}_1)\widehat{C}_1\end{pmatrix} \tag{S.40}$$

has strictly negative entries, (where we have used the definitions for q_{ij} in (S.36)) hence the profit functions are strictly concave (in this Case 3), and hence there is a unique maximum. From (S.36), here $\tilde{\mathbf{q}}$ is given by:

$$\tilde{\mathbf{q}} = \widehat{C}_2\begin{pmatrix}-2C_1&\widehat{C}_1&-\widehat{C}_1+b-1\\\widehat{C}_1&-2\widehat{C}_1&\widehat{C}_1-b+1\\-\widehat{C}_1&\widehat{C}_1&-2(\frac{1}{r}+1+\widehat{C}_1)\end{pmatrix}$$

and hence $\det\tilde{\mathbf{q}} = -2\widehat{C}_1^5 r^2[3+r(2+b+2\widehat{C}_1)] < 0$ (recall $\widehat{C}_2 = r\widehat{C}_1$). The left hand side of (S.38) is

$$\begin{pmatrix}-tq_{13}-c_1-c_3\\-c_2\\tq_{33}-c_3\end{pmatrix} = -\widehat{C}_1\begin{pmatrix}1-\widehat{C}_2 t\\br\\1+(1+r+\widehat{C}_2)t\end{pmatrix}.$$

Using Cramer's rule

$$p_1^* = \frac{r\widehat{C}_1^2}{\det\tilde{\mathbf{q}}}\det\begin{pmatrix}\widehat{C}_1 rt&\widehat{C}_1&-[\widehat{C}_1-(b-1)]r\\-(br+2)\widehat{C}_1&-2\widehat{C}_1&[\widehat{C}_1-(b-1)]r\\-(1+r+r\widehat{C}_2)t\widehat{C}_1&\widehat{C}_1&-2(1+r+\widehat{C}_2)\end{pmatrix}.$$

Simplifying gives

$$p_1^* = \frac{1}{2\widehat{C}_1 r[3+r(2+b+2\widehat{C}_1)]}\Big((2+br)[2+br+r(1+\widehat{C}_1)] + rt\Big[(1+r)(b-1)-\widehat{C}_1\,[5(1+r)+4r\widehat{C}_1]\Big]\Big).$$

Hence if $t=0$, it follows that p_1^* will be positive. Similarly,

$$p_2^* = \frac{1}{2\widehat{C}_1 r[3+r(2+b+2\widehat{C}_1)]}\Big(2+2(1+2b+\widehat{C}_1)r + b(3+b+3\widehat{C}_1)r^2 - tr\Big[b-1+\widehat{C}_1+(b-1-\widehat{C}_1+2b\widehat{C}_1)r\Big]\Big) \tag{S.41}$$

Hence

$$p_2^* > 0 \Leftrightarrow 0 \le t < \frac{2+2(1+2b+\widehat{C}_1)r+b(3+b+3\widehat{C}_1)r^2}{r(b-1+\widehat{C}_1+(b-1-\widehat{C}_1+2b\widehat{C}_1)r)}.$$

Similarly,

$$p_3^* = \frac{2+br+(3+3r+4\widehat{C}_1 r)t}{6+2r(2+b+2\widehat{C}_1)} \tag{S.42}$$

and hence $p_3^* > 0$ for all $t \ge 0$.

D. Proof of Theorem III.9 and Corollary III.10

The following Lemma proves parts 1 and 2 of Theorem III.9. The candidate solution $\{p_i^*\}$ is a local optimum for each i. The requirements that the p^* induce a feasible solutions result in the condition (26) together with the requirement that $p_1^* \ge 0$. If these conditions are satisfied, then either the p_i^* constitute a Nash equilibrium, or, they are such that either network 1 or 2 could improve their profits by deviating, in which case p_i^* is an ϵ-equilibrium. In Lemma S.2 we prove part 3 of Theorem III.9, and also characterize the ϵ of part 2. Note that we first prove the theorems for general transfer price t, which includes the special case $t=0$ (c.f. Section F of this Supplementary Material).

1) Proof of part 1 and 2 of Theorem III.9:

Lemma S.1. *If a Nash equilibrium exists with positive prices $\{p_i^*\}$, given by* (20), (21), (22), *with both networks 1 and 2 and the Content Provider having users and each making positive profit, then the transfer price satisfies*

$$-\frac{2+br}{3+r(3+4\widehat{C}_1)} < t < \frac{2+br}{3+r+2br}. \tag{S.43}$$

Conversely, when (26) *is satisfied and the expression for p_1^* in* (20) *is positive, then the $\{p_i^*\}$ given by* (20), (21), (22) *constitute an ϵ-equilibrium where $\epsilon \ge 0$. Further, all the prices $\{p_i^*\}$ are positive, and networks 1 and 2 and the Content Provider all have users.*

Proof. Since the prices $\{p_i^*\}$ are a local optimum for each i, it follows that the p_i^* will be a non-degenerate $\epsilon-$equilibrium if and only if the prices are positive, the market is covered ($Q_1+Q_2=1$), the user masses are positive ($Q_{11}, Q_{13}, Q_2 > 0$), and the profits are positive. Since we solve the equations for the Q_i ensuring the constraint $Q_{11}+Q_{13}+Q_2 = 1$ is met, necessary and sufficient conditions are that each $Q_{11}^*, Q_{13}^*, Q_2^*$ is in $(0,1)$, $p_i* > 0$ and $\pi_i^* > 0$.

i) Since $Q_{11}^* = p_3^*$, the condition $Q_{11}^* \in (0,1)$ is equivalent to,

$$0 < p_3^* < 1 \tag{S.44}$$

ii) Using (23), $Q_{13}^* \in (0,1)$ is equivalent to

$$0 < p_3^* - t < \frac{1+br}{1+r+r\widehat{C}_1} \tag{S.45}$$

iii) By construction $Q_2^* + Q_1^* = 1$, hence the requirement $Q_2^* \in (0,1)$ is equivalent to requiring $Q_1^* \in (0,1)$, which from (23) is equivalent to

$$0 < p_1^* + t < \frac{1+br}{r\widehat{C}_1} \tag{S.46}$$

These three conditions, together with the requirement that $p_1^* \ge 0$ also ensure that each $p_i^* \ge 0$ and each $\pi_i^* \ge 0$. Using inequality (S.44) and substituting from (S.42) gives the condition

$$-\frac{2+br}{3+r(3+4\widehat{C}_1)} < t < \frac{4+r(4+b+4\widehat{C}_1)}{3+r(3+4\widehat{C}_1)} \tag{S.47}$$

Using expression (S.45) and substituting from (S.42) gives the condition

$$-\frac{4+r(2+7b+2\widehat{C}_1+b(3+2b+3\widehat{C}_1)r)}{(3+r+2br)(1+r+\widehat{C}_1 r)} \;<\; t \;<\; \frac{2+br}{3+r+2br} \tag{S.48}$$

The conjunction of (S.47) and (S.48) gives the condition

$$-\frac{2+br}{3+r(3+4\widehat{C}_1)} \;<\; t \;<\; \frac{2+br}{3+r+2br}.$$

This condition also ensures that (S.46) is satisfied, completing the proof of the lemma. □

2) Proof of part 3 of Theorem III.9, and characterization of ϵ of part 2.:

Lemma S.2. *A Nash equilibrium exists if t satisfies* (S.43) *and in addition:*

$$\Big(\Big[(2+br)(2+br+r(1+\widehat{C}_1))\Big]+ \tag{S.49}$$

$$rt\Big[(1+r)(b-1)-\widehat{C}_1\Big(5(1+r)+4r\widehat{C}_1\Big)\Big]\geq 0 \tag{S.50}$$

or $t\leq 0\Big)$ *[condition for p_1^* to be positive]*

$$\text{and}\;\left(t\geq\frac{2+br}{1+r(1+\widehat{C}_1)}\quad\text{or}\right. \tag{S.51}$$

$$t\leq\frac{(2+br)(2+(1+\widehat{C}_1+b)r)}{6+r(11+b+15\widehat{C}_1+(b+2b\widehat{C}_1+(1+\widehat{C}_1)(5+8\widehat{C}_1))r)} \tag{S.52}$$

or t satisfies expression (S.56)$\Big)$ *[for p_1^* to be optimal]*

$$\text{and}\;\left(t\leq-\frac{1+r}{br}\qquad\text{or}\right. \tag{S.53}$$

$$t\leq\frac{4+r\left(-b^2r-b[2+r(1+\widehat{C}_1)]+4(1+\widehat{C}_1)[2+r(1+\widehat{C}_1)]\right)}{6+r(11+b+9\widehat{C}_1+(b+(1+\widehat{C}_1)(5+4\widehat{C}_1))r)} \tag{S.54}$$

or t satisfies expression (S.58) $\Big)$ *[cond. for p_2^* to be optimal]*

Proof. We prove that p^* is a Nash equilibrium by fixing two of $\{p_1^*,p_2^*,p_3^*\}$ while allowing the other p_i to vary, then showing conditions under which p_i^* is optimal for π_i.

p_1^* is optimal for network 1. With $p_2=p_2^*$, $p_3=p_3^*$, as we *increase* p_1 from p_1^* we either stay in Region 3, or potentially move into Region 1. There are three mutually exclusive cases we need to consider:

i) For all $p_1\geq p_1^*$ we remain in Region 3 and never move to Region 1, and hence p_1^* is optimal.
The boundary between Regions 3 and 1 occurs when, from (S.23) and (S.10), $p_1=p_1^B$ solves $p_3^*=[1+r(p_2^*-p_1)\widehat{C}_1]/[1+r(1+\widehat{C}_1)]$, (c.f. (S.73)), which substituting gives $p_1^B=[2+br-[1+r(1+2\widehat{C}_1)]t]/2r\widehat{C}_1$. Hence p_1 will stay in Region 3 if the boundary point is infeasible, $p_1^B\leq 0$, that is, if $t\geq\frac{2+br}{1+r(1+\widehat{C}_1)}$, i.e., (S.51).

ii) $\pi_i(p_1)$ is decreasing in Region 1 and hence $\pi_1(p_1)<\pi_1(p_1^*)$ $\forall p_1\in$ Region 1.
Since $\pi_1(p_1)$ is convex in Region 1, a sufficient condition for this is $\frac{\partial\pi_1}{\partial p_1}|_{p_1=p_1^B}\leq 0$, which substituting and taking derivatives in (S.6), using (S.5) and substituting $p_2=p_2^*$, $p_3=p_3^*$, $p_1=p_1=p_1^B$ gives (S.52).

iii) There is a feasible local maximum for π_1 in Region 1, where the profit is given by $\breve{\pi}_1=\breve{p}_1\breve{Q}_1$, but

$$\pi_1^*\;=\;p_1^*Q_1^*\;+\;tQ_{13}^*\;\geq\;\breve{\pi}_1, \tag{S.55}$$

and hence again p_1^* is optimal for π_1. The point $\breve{p}_1$ is where $\frac{\partial\pi_1}{\partial p_1}=0$, that is when $0=\frac{1+r(p_2^*-\breve{p}_1)\widehat{C}_1}{1+r(1+\widehat{C}_1)}-\breve{p}_1\frac{r\widehat{C}_1}{1+r(1+\widehat{C}_1)}$ and hence $\breve{p}_1=\frac{1+r\widehat{C}_1p_2^*}{2r\widehat{C}_1}$. At this point, the profit is given by $\breve{\pi}_1=\breve{p}_1\breve{Q}_1$, which is

$$\breve{\pi}_1\;=\;\breve{p}_1^2\frac{r\widehat{C}_1}{1+r(1+\widehat{C}_1)}\;=\;\frac{(1+r\widehat{C}_1p_2^*)^2}{4r\widehat{C}_1(1+r(1+\widehat{C}_1))}.$$

Substituting for Q_1^* from (23) in (S.55) gives the full condition as following the quadratic relation on t,

$$p_1^*\frac{r\widehat{C}_1}{1+br}(p_1^*+t)+t\frac{1+r(1+\widehat{C}_1)}{1+br}(p_3^*-t)\;\geq\;\frac{(1+r\widehat{C}_1p_2^*)^2}{4r\widehat{C}_1(1+r(1+\widehat{C}_1))} \tag{S.56}$$

where p_i^* are given in (20),(21),(22).
In the case that (S.56) does not hold (which necessarily also requires that (S.51) and (S.52) are not satisfied), define

$$\epsilon_1\;:=\;\frac{(1+r\widehat{C}_1p_2^*)^2}{4r\widehat{C}_1(1+r(1+\widehat{C}_1))}-p_1^*\frac{r\widehat{C}_1}{1+br}(p_1^*+t)+t\frac{1+r(1+\widehat{C}_1)}{1+br}(p_3^*-t).$$

Finally, if we *decrease*, p_1, we potentially move to Region 2. But we know that π_1 is decreasing in Region 2, and hence $\pi_1(p_1)\leq\pi(p_1^{B23})<\pi_1(p_1^*)$ f or all $p_1\in$ Region 2 where p_1^{B23} is the value of p_1 at the boundary of Regions 2 and 3.

p_2^* is optimal for network 2. The proof mirrors the arguments for showing p_1^* is optimal for network 1. With $p_1=p_1^*$, $p_3=p_3^*$, as we *decrease* p_3 from p_2^* we either stay in Region 3, or move into Region 1. There are three mutually exclusive cases we need to consider:

i) For $p_2\leq p_2^*$ we remain in Region 3 and never move to Region 1.
At the boundary point, p_2^B solves $p_3^*=\frac{1+r(p_2-p_1^*)\widehat{C}_1}{1+r(1+\widehat{C}_1)}$, that is, $p_2^B=\frac{t+r(b+t)}{2\widehat{C}_1r}$. The condition that is infeasible ($p_2^B<0$) or zero gives condition (S.53)

ii) $\pi_2(p_2)$ is increasing Region 1 and hence $\pi_2(p_2)<\pi_2(p_2^*)$ $\forall p_2\in$ Region 1, since $\pi_2(p_2)$ is concave in the interior of Region 1 .
Now using (S.7), differentiating and substituting $p_1=p_1^*,p_2=p_2^B,p_3=p_3^*$ gives that $\frac{\partial\pi_2}{\partial p_2}|_{p_2^B}=1-p_3^*-\frac{t+r(b+t)}{2[1+r(1+\widehat{C}_1)]}$ which will be non-negative if and only if (S.54) holds.

iii) There is a feasible local maximum for π_2 in Region 1 at the point $\breve{p}_2$, with profit given by by $\breve{p}_2\breve{Q}_2$, but for which

$$\breve{\pi}_2\;\leq\;\pi_2^*=p_2^*Q_2^* \tag{S.57}$$

and hence p_2^* is optimal for π_2. $\breve{p}_2$ is the point in Region 1 at which $\frac{\partial\pi_2}{\partial p_2}=0$, which using (S.7) and (S.5) gives the point $\breve{p}_2=\frac{1+\widehat{C}_1+\widehat{C}_1p_1^*}{2\widehat{C}_1}$. Using (20) gives the profit

$$\breve{\pi}_2=\left(\frac{1+\widehat{C}_1+\widehat{C}_1p_1^*}{2\widehat{C}_1}\right)^2\frac{r\widehat{C}_1}{1+r(1+\widehat{C}_1)}.$$

Q_2^* is given by (21), and (23) and hence substituting (S.57) is the condition

$$\left(\frac{1+\widehat{C}_1+\widehat{C}_1p_1^*}{2\widehat{C}_1}\right)^2\frac{r\widehat{C}_1}{1+r(1+\widehat{C}_1)}\leq p_2^*\frac{r\widehat{C}_1}{1+br}p_2^* \tag{S.58}$$

a quadratic in t, where the p_i^* are given in (20),(21).
In the case that (S.58), (S.53), (S.54) all fail to hold, define

$$\epsilon_2\;:=\;p_2^*\frac{r\widehat{C}_1}{1+br}p_2^*-\left(\frac{1+\widehat{C}_1+\widehat{C}_1p_1^*}{2\widehat{C}_1}\right)^2\frac{r\widehat{C}_1}{1+r(1+\widehat{C}_1)}.$$

Finally, if network 2 *increases* its price above p_2^*, it potentially moves to region 2; but network 2 receives zero profit in Region 2, hence network 2 has no incentive to increase its price above p_2^*.

p_3^* is optimal for the Content Provider. We show that under the conditions of the lemma, p_3^* is optimal for the Content Provider with no further restrictions.

i) If $\widehat{C}_1 \geq b-1$: As we decrease p_3, we remain in Region 3 and hence p_3^* is optimal for all $p_3 \leq p_3^*$. Moving to Region 2 is not possible since the boundary is infeasible: the Region 2-3 boundary is the point $p_3 \geq t$ which satisfies $p_3^B = \frac{(p_2^*-p_1^*)-b}{1-b+\widehat{C}_1}$. Substituting for p_1^* and p_2^* gives

$$\begin{aligned} p_3^B \;=\; & \Big(-1-br(3+(1+\widehat{C}_1+b)r) \\ & +(2\widehat{C}_1+1-b)r[1+r(1+\widehat{C}_1)]t\Big) \\ & \div\Big((\widehat{C}_1+1-b)r[3+(2+b+2\widehat{C}_1)r]\Big) \end{aligned} \tag{S.59}$$

and when $-\frac{2+br}{3+3r+4r\widehat{C}_1} \leq t \leq \frac{2+br}{3+r+2br}$, this implies $p_3^B < t$, and hence we never move to Region 2. Conversely, *increasing* p_3 either causes us to remain in Region 3, or move potentially move to Region 1 where the Content Provider receives zero profit, hence p_3^* is optimal for all $p_3 > p_3^*$.

ii) $\widehat{C}_1 < b-1$: the only way to violate (S.24) is to *increase* p_3. In this scenario with p_1^* and p_2^* then (S.59) with the condition $0 \leq t \leq \frac{2+br}{3+r+2br}$ implies $p_3 \geq 1$ and hence $\pi_3 = 0$. If instead $-\frac{2+br}{3+3r+4r\widehat{C}_1} \leq t < 0$, then to have an interior maximum in Region 2, it is necessary for both $p_3^B < 1$ and $p_3 < (1+t)/2$, and we can show that these three conditions cannot simultaneously hold, and hence if we enter Region 2, the value of π_3 will decrease. Hence p_3^* is optimal .

Summary. Necessary and sufficient conditions for a non-degenerate Nash equilibrium to exist at p^* in are that Lemma S.2 holds, i.e., (S.43) and {(S.50) or $t \leq 0$} and {(S.51) or (S.52) or (S.56)} and { (S.53) or (S.54) or (S.58)} hold.

When only the necessary conditions hold ((S.43) and {(S.50) or $t \leq 0$}) but not all the other conditions for Lemma S.2, (so either {(S.51) and (S.52) and (S.56)} are all false, or {(S.53) and (S.54) and (S.58)} are all false,) then p^* is an ϵ-equilibrium, not a Nash equilibrium, where $\epsilon = \max\{\epsilon_1, \epsilon_2\}$. ☐

3) Proof of Corollary III.10:

Proof. From (20) and the conditions of the corollary, we obtain $p_1^* > 0$. The result follows from part 2 of Theorem III.9. ☐

E. Proof of Theorem III.11

We provide here a more detailed presentation of Theorem III.11 than is given in Section III.

THEOREM III.11. There are only three possibilities for degenerate equilibria. Specifically, there exists a value t^A (which can be computed) such that:

1) If $t \geq t^A$, then there exists a Nash equilibrium in which the Content Provider prices itself out of the market by setting $p_3^* = t$, $Q_{13}^* = 0$, and for networks 1 and 2, the prices, user masses, and profits are given by

$$p_1{}^* \;=\; \frac{2+r+r\widehat{C}_1}{3r\widehat{C}_1}, \qquad p_2{}^* \;=\; \frac{1+2r+2r\widehat{C}_1}{3r\widehat{C}_1} \tag{S.60}$$

where the user masses are

$$Q_1^* \;=\; \frac{2+r+r\widehat{C}_1}{3(1+r+r\widehat{C}_1)} \qquad Q_2^* \;=\; \frac{1+2r+2r\widehat{C}_1}{3(1+r+r\widehat{C}_1)} \tag{S.61}$$

and the profits are

$$\pi_1^* \;=\; \frac{(2+r+r\widehat{C}_1)^2}{9r\widehat{C}_1(1+r+r\widehat{C}_1)} \qquad \pi_2^* \;=\; \frac{[1+2(1+\widehat{C}_1)r]^2}{9r\widehat{C}_1(1+r+r\widehat{C}_1)}. \tag{S.62}$$

2) If $t < 0$, then network 1 provides a subsidy to the Content Provider for each user, i.e., let $s := -t$. Then if the subsidy s is sufficiently great, viz., it is at least $\frac{2+br}{3+r(3+4\widehat{C}_1)}$, then there will be a unique Nash equilibrium where the Content Provider sets $p_3^* = 0$, and $Q_{11}^* = 0$. There are two subcases:

a) If $s \leq \frac{2+br}{r\widehat{C}_1}$, then the equilibrium is:

$$p_1^* \;=\; \frac{2+r(b+2\widehat{C}_1 s)}{3\widehat{C}_1 r}, \qquad p_2^* \;=\; \frac{1+r(2b+\widehat{C}_1 s)}{3\widehat{C}_1 r} \tag{S.63}$$

where

$$Q_{13}^* \;=\; \frac{2+r(b-\widehat{C}_1 s)}{3(1+br)}, \qquad Q_2^* \;=\; \frac{1+r(2b+\widehat{C}_1 s)}{3(1+br)} \tag{S.64}$$

with profits given by

$$\begin{aligned} \pi_1 \;&=\; \frac{[2+r(b-\widehat{C}_1 s)]^2}{9r\widehat{C}_1(1+br)}, \\ \pi_2 \;&=\; \frac{[1+r(2b+\widehat{C}_1 s)]^2}{9r\widehat{C}_1(1+br)}, \\ \pi_3 \;&=\; \frac{[2+r(b-\widehat{C}_1)]s}{3(1+br)}. \end{aligned} \tag{S.65}$$

b) If $s > \frac{2+br}{r\widehat{C}_1}$, then network 1 chooses a price of at least s; in consequence all users will choose network 2, $Q_2^* = 1$, and the equilibrium is: $p_1^* = s$, $p_2^* = s - \frac{1}{r\widehat{C}_1}$, with network 2 capturing all the profit, $\pi_2^* = s - \frac{1}{r\widehat{C}_1}$.

3) If $-\frac{2+br}{3+r(3+4\widehat{C}_1)} < t < t^A$, there exists a set of parametric conditions under which a Nash equilibrium exists where the optimal strategy for network 1 is to set its price to zero. Specifically, these are:

$$\widehat{C}_1 > b + 1/r \tag{S.66}$$

$$\frac{(2+br)(2+br+(1+\widehat{C}_1)r)}{r\{(1+r)(1-b)+\widehat{C}_1\,[5(1+r)+4\widehat{C}_1 r]\}} \;\leq\; t \;<\; \frac{2+br}{2+br+(1+\widehat{C}_1)r} \tag{S.67}$$

$$\frac{[(2+br)(1-t)-(1+\widehat{C}_1)rt]t}{1+br} \;\geq\; \frac{[2(2+br)+(1+\widehat{C}_1-b)rt)]^2}{4\widehat{C}_1 r(4+[3(1+\widehat{C}_1)+b]\,r)}, \tag{S.68}$$

which necessarily imply $t > 0$. The unique equilibrium is:

$$p_1^* = 0 \tag{S.69}$$

$$p_2^* = \frac{1+b+\widehat{C}_1+2br(1+\widehat{C}_1)+(1+\widehat{C}_1-b)[1+r(1+\widehat{C}_1)]t}{\widehat{C}_1\,(4+[3(1+\widehat{C}_1)+b]\,r)} \tag{S.70}$$

$$p_3^* = \frac{2+br+2[1+r(1+\widehat{C}_1)]t}{4+[3(1+\widehat{C}_1)+b]\,r}, \tag{S.71}$$

where

$$Q_{11}^* = p_3^*, \quad Q_2^* = \frac{r\widehat{C}_1}{1+br}p_2^*, \quad Q_{13}^* = \frac{1+r(1+\widehat{C}_1)}{1+br}(p_3^*-t),$$

and where the profits π_1, π_2, and π_3 can be calculated from (7).

Proof. First note that no Nash equilibrium is possible in Region 2, where $\nabla_{12} > \nabla_{13}$, since in that case the feasibility condition is $p_2 > p_1 + p_3 + p_3 + b(1-p_3)/\widehat{C}_1$, where the r.h.s is strictly greater than zero, and thus network 2 can decrease its price until equality holds, attracting users and moving out of Region 2 into Region 3. We first consider the three parts of the theorem, and then show that no other degenerate Nash equilibria exist.

---

1) $\mathbf{t \geq t^A}$*:* If the transfer price t is sufficiently large, the condition $p_3 \geq t$ implies $Q_{13} = 0$, and hence we are in Case 1 of Lemma III.1.

From (A.9), we have in the non-degenerate case the optimal critical value of w, is given by:

$$\nabla_{12} \;=\; \frac{(p_2-p_1)r\widehat{C}_1+1}{r\widehat{C}_1+r+1}.$$

The profits are $\pi_1 = p_1\nabla_{12} = p_1\left[\frac{(p_2-p_1)r\widehat{C}_1+1}{r\widehat{C}_1+r+1}\right]$, and $\pi_2 = p_2(1-\nabla_{12}) = p_2\left[\frac{[1-(p_2-p_1)]r\widehat{C}_1+r}{r\widehat{C}_1+r+1}\right]$. From the first derivatives, we obtain the unique optimal prices (S.60), (optimal since the second derivatives are negative). The optimal critical value of w, i.e., the optimal value of ∇_{12}, is given by:

$$\nabla^*_{12} \;=\; \frac{2+r+r\widehat{C}_1}{3(1+r+r\widehat{C}_1)} \;=\; \frac{2+r+\widehat{C}_2}{3(1+r+\widehat{C}_2)}. \tag{S.72}$$

By (S.5), the mass of users on the networks at equilibrium are given by (S.61). □

The optimal choice of p_1, p_2, namely p_1^*, p_2^*, are given by (S.60) and the Q_i^* from (S.61). Since $Q_{13} = 0$ and $Q_{11} = Q_1$, we know that (S.60) and $p_3^* \geq t$ constitute a local equilibrium. To prove that these values constitute a Nash equilibrium, we need to show that for $t \geq t^A$, given p_2^* and $p_3{}^*$, network 1 cannot benefit by altering its price from p_1^*, with corresponding statements for network 2 and the Content Provider.

$\mathbf{p_3^* = t}$ **is optimal for the Content Provider**. At this value the Content Provider has zero profit. For the Content Provider to have a nonzero profit, we require $p_3 \geq t$, hence the Content Provider cannot lower its price below this value. With the given values of p_1^*, p_2^*, rasing the price above t also generates zero profit. Hence $p_3 = t$ is optimal for the Content Provider.

By Lemma III.1, we must have: $Q_2^*/r\widehat{C}_1 - gQ_1^*/\widehat{C}_1 < p_1^*+p_3-p_2^*$. Substituting in (27) implies this will hold iff p_3 is greater than $[2+r(1+\widehat{C}_1)]/[3(1+r(1+\widehat{C}_1))]$. Hence we must have

$$t^A \;>\; \frac{2+r(1+\widehat{C}_1)}{3[1+r(1+\widehat{C}_1)]},$$

since otherwise the Content Provider could lower its price, p_3, below this value to attract users until equality holds.

$\mathbf{p_1^*}$ **is optimal for network 1**.

For the given values p_1^* and p_2^*, we are in the middle subcase of the alternative defining condition for Case 1 of Lemma III.1, (S.10). If network 1 raises its price from p_1^*, it will remain in Case 1 and hence p_1^* will remain the optimal response to p_2^*.

If network 1 decreases its price from p_1^*, it is possible that $\pi_1(p_1; p_2^*; p_3^*) > \pi_1(p_1^*; p_2^*; p_3^*)$, in which case p_1^* is *not* an equilibrium. It is straightforward to shown that this can only happen if p_1 moves to be in Region 3, and has a greater local optimum in region 3 . We need to consider the cases (A) $p_3^* = t \leq 1$ and (B) $t > 1$ separately.

We consider subcase B first.

(B) $p_3^* > 1$. In this subcase, the first inequality in (S.10) is violated by becoming an equality as p_1 is decreased. By (S.31), $Q_1 = 1$, and by (S.32), at this point $p_1 = p_2^* - p_3^* - \frac{1}{\widehat{C}_1}$, with profit π_1 less than the profit at p_1^*, i.e.

$$\begin{aligned}\pi_1 \;=\; p_1 Q_1 \;=\; p_1 \;&<\; p_2^* - 1 - \frac{1}{\widehat{C}_1}\\ &=\; \frac{1-r-rc}{3r\widehat{C}_1}\\ &<\; \frac{(2+r+r\widehat{C}_1)^2}{9r\widehat{C}_1[1+r(1+\widehat{C}_1)]} \;=\; p_1^* Q_1^*.\end{aligned}$$

If we were to reduce p_1 even further, then we would immediately move to Case 2 Theorem III.1, since (S.32) is an equality, where $\pi_1(p_1)$ decreasing as we decrease p_1. Hence no higher value of the profit is possible in case B.

(A) $p_3^* \leq 1$. For this to have a local maximum in Region 3 such that $\pi_1(p_1; p_2^*; p_3^*) > \pi_1(p_1^*; p_2^*; p_3^*)$ we require:

(i) The boundary between Region 1 and Region 3 to be feasible for p_1. At this boundary point the second inequality in (S.10) is violated by becoming an equality,

$$t = p_3^* \;=\; \frac{1+r(p_2^*-p_1^{13})\widehat{C}_1}{1+r(1+\widehat{C}_1)}. \tag{S.73}$$

Substituting for p_2^* from (27) and simplifying gives the condition that $p_1^{13} > 0$ if and only if

$$t < \frac{4+2r+2r\widehat{C}_1}{3+3r+3r\widehat{C}_1}. \tag{S.74}$$

(ii) The derivative of the profit $\partial\pi_1/\partial p_1$ at the boundary p_1^{13} is negative. It follows from (7), (S.21), (S.19), (S.18), and from $Q_{13} = 0$, and (S.73), that:

$$\frac{\partial\pi_1}{\partial p_1}|_{p_1^{13}} \;=\; p_1\frac{\partial Q_1}{\partial p_1} + Q_1 + t\frac{\partial Q_{13}}{\partial p_1} \;=\; t - (p_1^{13}+t)\frac{r\widehat{C}_1}{1+br}.$$

Substituting for p_1 from (S.73) and for p_2^* from (27) yields the condition:

$$t \;<\; \frac{2(2+r+r\widehat{C}_1)}{3(2+r+br)} \quad\Longleftrightarrow\quad \frac{\partial\pi_1}{\partial p_1} \;<\; 0. \tag{S.75}$$

(iii) There is a feasible local optimum in region 3. Taking derivatives $\partial\pi_1/\partial p_1$ using (7), (S.21), (S.19), (S.18), and solving for p_1^0 such that $\partial_{p_1}\pi_1|p_1^0 = 0$, gives on substituting $p_3^* = t$,

$$p_1^0 = \frac{1+p_2^* r\widehat{C}_1 + (b-1-2\widehat{C}_1)rt}{2r\widehat{C}_1}$$

and hence substituting for p_2^*

$$t(3r[1-b+2\widehat{C}_1]) < 2(2+2r+2r\widehat{C}_1) \quad\Longleftrightarrow\quad p_1^0 > 0. \tag{S.76}$$

(iv) The feasible local optimum generates higher profit. That is $\pi_1(p_1^0; p_2^*; p_3^*) > \pi_1(p_1^*; p_2^*; p_3^*)$. Substituting gives the condition

$$\begin{aligned}&\frac{1}{36r\widehat{C}_1(1+br)}\Big(4(2+r+r\widehat{C}_1)^2+\\ &12(b-1)r(2+r+r\widehat{C}_1)t + 9r\left((b-1)^2r - 4c(1+br)\right)t^2\Big)\\ &\qquad > \frac{(2+r+r\widehat{C}_1)^2}{9r\widehat{C}_1(1+r+r\widehat{C}_1)}\end{aligned}$$

which is equivalent to the condition

$$\begin{cases} t < t_l(b,r,\widehat{C}_1) \text{ OR } t > t_u(b,r,\widehat{C}_1) & \text{if } (b-1)^2r \geq 4\widehat{C}_1(1+br)\\ t_l(b,r,\widehat{C}_1) < t < t_u(b,r,\widehat{C}_1) & \text{if } (b-1)^2r < 4\widehat{C}_1(1+br)\end{cases} \tag{S.77}$$

where t_l, t_u are the upper and lower roots of the equation

$$9\left((b-1)^2 r - 4\widehat{C}_1(1+br)\right)t^2 + 12(b-1)(2+r(1+\widehat{C}_1))t - \frac{4}{1+r(1+\widehat{C}_1)}(b-1-c)(2+r(1+\widehat{C}_1))^2 = 0.$$

With a slight abuse of notation, we shall let (S.74) etc to refer to the conditions on t: hence p_1^* is *not* optimal for network 1 only if (S.74) AND (S.75) AND (S.76) AND (S.77) hold, thus p_1^* is optimal if NOT ((S.74) AND (S.75) AND (S.76) AND (S.77)).

$\mathbf{p_2^*}$ is optimal for network 2.
The proof mirrors that for showing p_1^* is optimal. For network 2, if we decrease p_2 from p_2^*, we remain in Case 1 of Theorem III.1, and hence cannot improve upon $\pi_2(p_2^*)$.

As we increase p_2 above p_2^* it is possible $\pi_2(p_1^*;p_2;p_3^*) > \pi_2(p_1^*;p_2^*;p_3^*)$, in which case p_2^* is *not* an equilibrium. For this to happen, p_2 must force a move to region 3, and network 2 must have a greater local optimum in that region. We need to consider the cases (A) $p_3^* = t \le 1$ and (B) $t > 1$ separately.

For case (B), $p_3^* = p_2 - \frac{1}{\widehat{C}_1} + p_1^*$. At boundary point $\pi_2 = 0$ and as we increase π_2 still further we move into Case 2, hence no greater profit for network 2 is possible in this case.

(A) $\pi_2(p_1^*;p_2;p_3^*) > \pi_2(p_1^*;p_2^*;p_3^*)$ for some p_2 requires

(i) The boundary between Region 1 and Region 3 to be feasible for p_2. At the boundary $p_2 = p_2^{13}$ solves

$$t = p_3^* \;=\; \frac{1 + r(p_2^{13} - p_1^*)\widehat{C}_1}{1 + r + r\widehat{C}_1}. \tag{S.78}$$

Substituting p_1^* from (27), and simplifying gives the condition that p_2^{13} will be positive is

$$t > \frac{1 - r - \widehat{C}_1 r}{3 + 3r + 3\widehat{C}_1 r}. \tag{S.79}$$

(ii) The derivative $\partial\pi_2/\partial p_2$ at the boundary $p_2 = p_2^{13}$ is positive. Now

$$\frac{\partial \pi_2}{\partial p_2} \;=\; -p_2 \frac{r\widehat{C}_1}{1+br} + (1 - p_3^*).$$

Using (S.78) and (27), this will be positive if

$$t = p_3^* < \frac{4 + 3br - (1+\widehat{C}_1)r}{3[2 + br + (1+\widehat{C}_1)r]}. \tag{S.80}$$

(iii) There is a feasible local optimum in region 3. The local optimum for π_2 in region occurs at the price p_2^0 where $\partial_{p_2}\pi_2|p_2^0 = 0$, which substituting gives the value

$$p_2^0 = \frac{2 + r + 3br + r\widehat{C}_1 + 3(1 - b + \widehat{C}_1)rt}{6r\widehat{C}_1}$$

which will be feasible provided that

$$2 + r + 3br + r\widehat{C}_1 + 3(1 - b + \widehat{C}_1)rt > 0. \tag{S.81}$$

(iv) The feasible local optimum generates higher profit. The profit at p_2^0 is given by

$$\pi_2(p_1^*;p_2^0;p_3^*) = \frac{(2 + r + 3br + r\widehat{C}_1 + 3(1 - b + \widehat{C}_1)rt)^2}{36r\widehat{C}_1(1+br)}$$

and this will be greater than $\pi_2(p_1^*;p_2^*;p_3^*) = \frac{(1+2(1+\widehat{C}_1)r)^2}{9r\widehat{C}_1(1+r+r\widehat{C}_1)}$ provided that

$$t < t_l(b,r,\widehat{C}_1) \text{ OR } t > t_u(b,r,\widehat{C}_1) \tag{S.82}$$

where t_l, t_u are the upper and lower roots of the quadratic in t

$$(2 + r + 3br + r\widehat{C}_1 + 3(1 - b + \widehat{C}_1)rt)^2 = \frac{4(1+br)(1+2(1+\widehat{C}_1)r)^2}{1+r+rc}.$$

Hence p_2^* is *not* optimal only if (S.79) AND (S.80) AND (S.81) AND (S.82) hold, and hence p_2^* is optimal if NOT ((S.79) AND (S.80) AND (S.81) AND (S.82)).

$\mathbf{p_3^*}$ is optimal for the Content Provider. Trivial.

Summarizing, hence necessary and sufficient conditions for a Nash equilibrium to exist in this case are that NOT ((S.79) AND (S.80) AND (S.81) AND (S.82)) AND NOT ((S.74) AND (S.75) AND (S.76) AND (S.77)). Hence by choosing the simpler conditions in this expression, it follows that *sufficient* conditions for a Nash equilibrium to exist with these p_i^* are that $t \ge t^A$, where

$$t^A = \min\Bigg\{1, \max\left(\frac{4 + 3br - (1+\widehat{C}_1)r}{3[2 + br + (1+\widehat{C}_1)r]}, \frac{2 + r(1+\widehat{C}_1)}{3[1 + r(1+\widehat{C}_1)]}, \frac{2[2 + r(1+\widehat{C}_1)]}{3[2 + br + r]}\right), \max\left(\frac{4 + 3br - (1+\widehat{C}_1)r}{3[2 + br + (1+\widehat{C}_1)r]}, \frac{2[2 + r(1+\widehat{C}_1)]}{3[1 + r(1+\widehat{C}_1)]}\right)\Bigg\}. \tag{S.83}$$

---

2) $\mathbf{t < 0,\ \ s = -t \ge \frac{2+br}{3+r(3+4\widehat{C}_1)}}$: e

When the subsidy is large enough, the Content Provider can set its price p_3 to zero. Now when $p_3 = 0$, then by (S.16), $Q_{11} = 0$, and basic service is never used in this case. First consider the case

$$\frac{2+br}{r\widehat{C}_1} \;\ge\; s \ge \frac{2+br}{3 + r(3+4\widehat{C}_1)}$$

corresponding to subcase 2(a) of Theorem III.5.

Proof. Proof of subcase 2(a) of III.11 It is straightforward to check that the p_i in the system (30) with the corresponding Q's are given in (31) are consistent with being in Region 3, and moreover p_1^*, p_2^* satisfy the first order conditions for π_1, π_2, i.e., $\frac{\partial \pi_i}{\partial p_i} = 0$, when $p_3 = 0$. The second order conditions are also satisfied for networks 1 and 2. The condition for π_3 to have a maximum at $p_3 = 0$ is that $\frac{\partial_{p_3}}{\partial \pi_3} \le 0$. For this to hold when $p = p^*$ requires:

$$\left.\frac{\partial \pi_3}{\partial p_3}\right|_{p=p^*} \;=\; s\frac{\partial Q_{13}}{\partial p_3} + Q_{13} \;=\; \frac{2 + br - (3 + 3r + 4\widehat{C}_1 r)s}{3 + 3br} \;\le\; 0 \tag{S.84}$$

and hence

$$s \;\ge\; \frac{2+br}{3 + r(3+4\widehat{C}_1)}. \tag{S.85}$$

The requirement that $Q_{13} \ge 0$ necessitates that

$$s \;\le\; \frac{2+br}{r\widehat{C}_1}.$$

When both conditions, (S.84) and (S.85), on s are satisfied, the remaining feasibility requirements ($Q_{13} \le 1$, $\pi \ge 0$) are also satisfied, hence we have shown that vector (p_i^*) is a local maximum. It remains to prove that these are globally optimum prices.

p_1^* *is optimal*. While we remain in Region 3, we know p_1^* is optimal. By increasing p_1, we remain in Region 3 until we reach the boundary with region 1, at which point from (S.19), $Q_{13} = 0$ (recall $p_3^* = 0$), hence $Q_1 = 0$, $\pi_1 = 0$ and this profit remains zero remains so as we increase p_1 further (c.f. (S.5) final condition). Now consider what happens when we decrease p_1 below p_1^*. If we decrease p_1 sufficiently, then we potentially enter Region 2. For this to happen, since $p_2 = p_2^*$, it follows from (S.15) that

$$p_1 \;<\; p_2^* - \frac{b}{\widehat{C}_1} \;=\; \frac{1 - br + r\widehat{C}_1 s}{3r\widehat{C}_1}.$$

But then, to have a higher profit than that corresponding to the prices given in (30), (30) we require

$$\frac{1 - br - 2r\widehat{C}_1 s}{3r\widehat{C}_1} \;>\; \frac{(2 + br - r\widehat{C}_1 s)^2}{9r\widehat{C}_1(1+br)}$$

which is clearly impossible for $s > 0$ (recall that $b > 1, r > 0, c > 0$). Hence p_1^* is optimal.

p_2^* *is optimal.* As we increase p_2 and remain in Region 3 (holding $p_3 = 0$ and $p_1 = p_1^*$), the profit for network 2 remains suboptimal (less than π_2^*), until we enter Region 2. But network 2 has zero profit in Region 2. Hence network 2 cannot increase its profit by increasing its price. If network 2 were to decrease its price from p_2^*, then again its profit would be suboptimal while in Region 3. If it were to decrease its price further to enter Region 1; then at the boundary of region and 1 and 3, from (S.20) $Q_{13} = 0$ (recall $p_3^* = 0$) and hence Q_2=1. The profit at this point is less than π_2^*, and decreasing pi_2 further decreases the profit further.

p_3^* *is optimal.* Since $p_3^* = 0$, the Content Provider can only increase its price, which potentially takes it to Region 1 where the Content Provider has no users and zero profit. Hence the Content Provider cannot increase its profit by changing its price.

Hence we have shown that p^* is indeed a Nash equilibrium, thus proving the subcase 2(a). □

Now consider case 2(b) of the theorem where

$$s > \frac{2 + br}{r\widehat{C}_1}.$$

Proof. Proof of subcase 2(b) of Theorem III.11 With the p_i given by $p_1^* = s$, $p_2^* = s - \frac{1}{r\widehat{C}_1}$, $p_3^* = 0$, we are in Region 3 with $Q_2 = 1$.

We now show that this is indeed a Nash equilibrium. If the Content Provider were to lower its price it would become negative and potentially move the scenario to Region 2. However, $p_3 < 0$ violates Region 2 feasibility condition (S.13). If the Content Provider were to raise its price, this could potentially move the scenario to Region 1, but in that case the Content Provider still receives zero profit. Thus the Content Provider will not change its price.

Network 2 has no incentive to lower its price, since network 2 already has all the users, and this could only result in lowering its profit. network 2 has no incentive to raise its price either, since this would potentially move the scenario to Region 2, but in Region 2 we have $Q_2 = 0$, and so network 2 would have no profit. Thus network 2 will not change its price.

If network 1 were to decrease its price below s, this would potentially move the scenario to Region 2; but Region 2 feasibility condition (S.14) is $p_1 - s(1-p_3) \geq 0$, which reduces here to $p_1 - s \geq 0$, a contradiction, since $p_1 - s < 0$. If network 1 were to increase its price, this would potentially move the scenario to Region 1; but in this case, (S.10) and (S.5) imply $Q_1 = 0$, and network 1 would still get no profit. Thus network 1 will not benefit by changing its price. □

---

3) $\mathbf{t > 0}$:

Proof. When $p_1 = 0$ and we are in Region 3, then from (7), $\pi_1 = tQ_{13}$, and using (S.19) and (S.21)

$$\frac{\partial \pi_1}{\partial p_1} \;=\; Q_1 + t\frac{\partial Q_{13}}{\partial p_1} \;=\; Q_1 - t\frac{r\widehat{C}_1}{1+br}$$
$$= \frac{1 + p_3(b - 1 - r\widehat{C}_1) + r\widehat{C}_1(p_2 - t)}{1+br}.$$

For $p_1 = 0$ to be a Nash equilibrium for network 1, it must first be a local maximum within Region 3, which requires

$$\pi_1 \geq 0 \quad \text{and} \quad \frac{\partial \pi_1}{\partial p_1} < 0. \tag{S.86}$$

In addition p_2 and p_3 need to be local maxima, and hence solve

$$\frac{\partial \pi_2}{\partial p_2} = 0, \;\; \frac{\partial \pi_3}{\partial p_3} = 0 \quad \text{with } p_1 = 0. \tag{S.87}$$

Using equations (S.19), (S.20), then (S.87) becomes

$$b + (1 + \widehat{C}_1 - b)p_3 \;=\; 2\widehat{C}_1 p_2$$
$$2p_3(1 + r + r\widehat{C}_1) \;=\; 1 + t(1 + r + r\widehat{C}_1) + r\widehat{C}_1 p_2.$$

Solving these equations gives (S.70) and (S.71) which is a local optimum since the $\pi_i(p_i)$ are convex. Substituting into (S.86) and using algebraic manipulation gives the conditions (S.66) and (S.67) on $b, r, \widehat{C}_1, t$ in the proposition, which also enure that $p_2^* > 0$ and $p_3^* > 0$. The profit for network 1 is then

$$\frac{1 + r(1 + \widehat{C}_1)}{1 + br} t(p_3^* - t). \tag{S.88}$$

For this to be Nash equilibrium, it must a be global optimum: in the case of network 1, network 1 can increase its price, and move the solution into Region 1. If does so, it has an optimum response to network 2 setting its price to p_2^*, namely the price

$$p_1 \;=\; \frac{(1 + r + r\widehat{C}_1)(4 + 2br + (1 - b + \widehat{C}_1)rt)}{2r\widehat{C}_1(4 + [b + 3(1+\widehat{C}_1)]r)}.$$

The condition that the resulting profit for network 1 at this value (calculated from (S.7)) does not exceed (S.88) gives (S.68). □

---

4) *No other degenerate Nash equilibria exist:*

To complete the proof of the theorem, we need only show that there cannot be a Nash equilibrium under any of the remaining boundary conditions, all of which must be in Region 3. We will have a Nash equilibrium at the point $p^* = (p_1^*, p_2^*, p_3^*)$ provided that the point p^* is feasible and, for each i, $\pi_i(p_i)$ is maximized at p^*. But as we have seen, from (S.39), that the profit functions $\pi_i(p_i)$ are strictly concave, hence the only possible Nash equilibria in this case occur either at a unique interior point of the feasible region or at the boundaries of their support. The boundaries of the region are characterized in Corollary III.3 (the corresponding intervals for each $\pi_i(p_i)$ are generated by the intersection of the lines formed by fixing $p_j, j \neq i$ in the boundaries). The boundaries correspond to the hyperplanes $p_1 = 0$, $p_2 = 0$ $p_3 = 0$, and $Q_{11} = 0$, $Q_{13} = 0$, $Q_2 = 0$ (and since by construction $Q_{11} + Q_{13} + Q_2 = 1$ holds, we also consider the constraints $Q_2 = 1$). In addition, when $t \geq 0$ we have the boundary $p_3 = t$ corresponding to $\pi_3(p_3) = 0$ (note that $\pi_1 = 0, \pi_2 = 0$ are covered by other boundaries). When $t \leq 0$, there is the additional constraint boundary $\pi_1 = 0$.

Condition $\boldsymbol{\pi_1 = 0}$. No Nash equilibrium can exist in this case, since if $\pi_1 = p_1 Q_{11} + (p_1 + t)Q_{13} = 0$, we must have $p_1 + t < 0$ (the degenerate case $Q_{11} = 0 = Q_{13}$ is covered by subcases below), then from from (7)

$$\frac{\partial \pi_1}{\partial p_1} \;=\; Q_1 + (p_1 + t)\frac{\partial Q_{13}}{\partial p_1} > 0$$

and network 1 can increase its profit away from 0 by increasing p_1.

Condition $\mathbf{Q_{13} = 0}$. When $Q_{13} = 0$, then from (7)

$$\frac{\partial \pi_3}{\partial p_3} \;=\; p_3\frac{\partial Q_{13}}{\partial p_3} < 0$$

unless $p_3 = 0$. Hence we can decrease p_3, thereby increasing the Content Provider's profit away from zero, unless either $p_3 = 0$ or $t > 0$ and $p_3 = t$. We treat each of these special cases below.

Condition $\mathbf{p_2 = 0}$. Here, to maximize π_2 we must have $\frac{\partial \pi_2}{\partial p_2}\Big|_{p_2=0} \leq 0$. From (7), $\frac{\partial \pi_2}{\partial p_2}\Big|_{p_2=0} = Q_2$, hence $Q_2 = 0$. Then from (12) and (13), $p_1 + p_3 < 0$, which implies at least one price is negative, and hence there is no feasible Nash equilibrium.

Condition $\mathbf{Q_2 = 0}$. (7) When $Q_2 = 0$, $\pi_2 = 0$ then to be in case 3 requires that $p_2 > 0$, while from (7)

$$\frac{\partial \pi_2}{\partial p_2} = p_2 \frac{\partial Q_2}{\partial p_2} < 0$$

and hence network 2 can decrease its price and generate positive profit. Thus $Q_2 = 0$ cannot be a Nash equilibrium for network 2.

Condition $\mathbf{p_3 = t}$.
From (7), when $p_3 = t$, $\frac{\partial \pi_3}{\partial p_3} = Q_{13}$, and if $Q_{13} > 0$, then this cannot be a Nash equilibrium for the Content Provider, since it could increase its profit by increasing t. The only possibility is that both $p_3 = t$ and $Q_{13} = 0$, which is included as a particular special case of Theorem III.11 part 1.

F. Proof of Theorem III.13 and Corollary III.14

Proof. Theorem III.13 is a special case of Theorem III.9 where $t = 0$, proved using Lemmas S.1 and S.2. Now (S.43), which gives conditions for prices and profits to be non-negative, is automatically satisfied when $t = 0$. This leaves Lemma S.2; when $t = 0$, the conditions of Lemma S.2 simplify: the conditions for p_1^* to be positive and p_1^* to be optimal are satisfied, leaving the conditions for π_2 to be optimal, ((S.54) or (S.58)), which reduce to

$$br[2 + (1+b+\widehat{C}_1)r] \;\leq\; 4(1+r+\widehat{C}_1 r)^2 \quad \text{OR} \tag{S.89}$$

$$\frac{4(2 + r(2 + 2\widehat{C}_1 + b(4 + [b + 3(1+\widehat{C}_1)]r)))^2}{1 + br} \geq \frac{\left(4 + r\left(b^2 r + 4(1+\widehat{C}_1)[2 + r(1+\widehat{C}_1)] + b(4 + 3(1+\widehat{C}_1)r)\right)\right)^2}{1 + r + \widehat{C}_1 r}. \tag{S.90}$$

We roll Theorem III.13 and Corollary III.14 into the following Lemma.

Lemma S.3. *When $t = 0$, sufficient conditions for a Nash equilibrium to exist are*

$$b \leq 2(1+\widehat{C}_1) + \frac{1}{r}. \tag{S.91}$$

Sufficient conditions for a Nash equilibrium not to exist are

$$r > 1 \textit{ AND } b > 4(1+\widehat{C}_1) + \frac{1}{r} \tag{S.92}$$

Proof. Proof of Lemma The first condition (S.89) is clearly satisfied if $b \leq 1 + \widehat{C}_1$. By writing $b = 2 + k\widehat{C}_1$, expanding the second inequality, taking out a factor of $(b - 1 - \widehat{C}_1)$ and equating the coefficients of r^i in the remaining quartic to ensure that the resulting polynomial is always positive , then we can show that resulting inequality will always be satisfied provided $k \leq 2$. That is, if $b \leq 2(1+\widehat{C}_1)$, a Nash equilibrium will always exist. A more detailed line of reasoning will produce a broader sufficient condition:

$$b \leq 2(1+\widehat{C}_1) \quad \text{or} \quad r \leq \frac{1}{b - 2(1+\widehat{C}_1)}.$$

which can be combined into the single condition (S.91).

By substituting and simplifying, we can also show that when the condition (S.92) holds then neither (S.89) nor (S.90) are true, and hence (S.92) is a sufficient condition for a Nash equilibrium not to exist. □

G. Proof of Theorem IV.1

We provide here a slightly more detailed presentation of Theorem IV.1 than is given in Section IV.

THEOREM IV.1 There exists a unique Nash equilibrium in the two-stage game. This occurs in one of two mutually exclusive cases, where in each case network 1 sets a positive transfer price t^*. In the first case all three parties make a positive profit; the second case is degenerate, where the Content Provider is shut out of the market. Specifically, either:

1) All three parties make a positive profit. The prices, user masses, and profits for this case are given by Theorem III.5 with $t = t^B$or t^O, where:
   a) t^O solves the first-order profit maximization conditions. This is the value of t satisfying first order conditions, namely
   $$\frac{\partial}{\partial t}\pi_1(p_1^*(t), p_2^*(t), p_3^*(t)) = 0.$$
   This will yield an affine equation in t, with solution t^O. Here t^O is the value of t that maximizes a concave profit function and hence can in principle be found in a straightforward way by network 1.
   b) t^B is the point at which network 2 is indifferent between competing with the Content Provider or lowering its price to drive it out of the market. Here $t^B < t^O$, where t^B is the feasible solution to the equation in t derived from (S.58), that is the positive solution to
   $$\left(\frac{1 + \widehat{C}_1 + \widehat{C}_1 p_1^*(t)}{2\widehat{C}_1}\right)^2 \frac{1}{1 + r(1+\widehat{C}_1)} = \frac{1}{1+br}(p_2^*(t))^2 \tag{S.93}$$
   a quadratic in t, where the p_i^* are given in (20),(21). Note that from Theorem III.9, equation (26), we must have $t^B < \frac{2+br}{3+r+2br}$.

   Sufficient conditions for this case to exist are that $b^*(r, \widehat{C}_1) \leq b \leq 1 + \widehat{C}_1$ OR $(1 + \widehat{C}_1 < b \leq 2(1 + \widehat{C}_1) + 1/r$
2) The equilibrium is degenerate, with the two networks making positive profit, and the Content Provider shut-out of the market. The prices, user masses and profits are given in Theorem III.11,with $t^* \geq t^A$, where t^A is defined in Theorem III.11 and t_A is given explicitly in (S.83), and satisfies $t_A \geq 1$. There are two subcases:
   a) b is small, satisfying $1 \leq b \leq b^*(r, \widehat{C}_1)$ where b^* is the root of a cubic equation, with the bound $b^* < 1 + \widehat{C}_1$. In this case network 1 makes strictly greater profit than having a positive transfer price.
   b) b is large, for which sufficient conditions are $b > 4(1 + \widehat{C}_1) + 1/r$ and $r > 1$ (c.f. Corollary III.14). In this instance network 1 cannot attain the higher profits that could be gained by setting a positive transfer price.

Proof. Proof We consider two subcases separately: $b \leq 1 + \widehat{C}_1$ and $b > 1 + \widehat{C}_1$. The statements of the theorem follow by combining the results from the subcases.

Subcase: When $b \leq 1 + \widehat{C}_1$ Recall that in this case the profit for network 1 is smaller with $t = 0$ than without the Content Provider.

i) First note that on calculating the profits from (7) using the prices and user masses from Theorem III.5, we can see that $\pi_1(p_1^*(t), p_2^*(t), p_3^*(t))$ is quadratic in t. By direct calculation, it is straightforward to see that
$$\frac{\partial^2}{\partial t^2}\pi_1(p_1^*(t), p_2^*(t), p_3^*(t)) < 0$$
for $1 \leq b \leq 1 + \widehat{C}_1$ hence any local optimum satisfying the first order conditions will be a global optimum if staying within the feasible region of Case 3.

ii) The value of t satisfying first order conditions is

$$t^O = (br+2)(\widehat{C}_1 r + r + 1)\left(b^2 r + 2b - \widehat{C}_1(r+1) - r - 2\right)$$
$$\div\Big[16\widehat{C}_1^3 r^2(br+1) + \widehat{C}_1^2 r(r(4b((b+8)r+10) - r + 30) + 35)$$
$$+ 2\widehat{C}_1(r+1)\left(r([b(b+9)-1]r + 11b + 7) + 9\right)$$
$$-(b-1)^2 r(r+1)^2\Big] \quad \text{(S.94)}$$

iii) For $1 \le b < \sqrt{\frac{\widehat{C}_1 r^2 + \widehat{C}_1 r + r^2 + 2r + 1}{r^2}} - \frac{1}{r}$, then $\frac{\partial}{\partial t}\pi_1(p_1^*(t), p_2^*(t), p_3^*(t))|_{t=0} < 0$, so network 1 can indeed increase revenue by reducing the transfer price from zero, setting a negative transfer price, thereby effectively subsidizing the Content Provider. However, network 1 can do even better by *raising* the transfer price to such a level that the Content Provider is shut out of the market—for example, by setting $t = 1$. The latter follows by first showing that $\frac{\partial}{\partial t}\pi_1(p_1^*(t), p_2^*(t), p_3^*(t))|_{t=-\frac{-2+br}{3+r(3+4\widehat{C}_1)}} > 0$, hence the value of t_0 is a potential Nash equilibrium (c.f. Theorem 5.3)); second, proving that $\frac{\partial}{\partial b}\pi_1(p_1^*(t^O), p_2^*(t^O), p_3^*(t^O)) > 0$ in this region, i.e., profits for network 1 increase with b; and third, showing that when $b = \sqrt{\frac{\widehat{C}_1 r^2 + \widehat{C}_1 r + r^2 + 2r + 1}{r^2}} - \frac{1}{r}$, network 1 is better shutting the Content Provider out of the market, hence is it better offer adopting this policy for all b in this range.

iv) For $\sqrt{\frac{\widehat{C}_1 r^2 + \widehat{C}_1 r + r^2 + 2r + 1}{r^2}} - \frac{1}{r} < b \le \widehat{C}_1 + 1$, then the optimal choice of strategy depends upon the values of $b, \widehat{C}_1, r$, or on a value $b^* = b^*(r, \widehat{C}_1)$, where b^* is the root of a cubic equation involving r and $\widehat{C}_1$. This is the value of b at which network 1 is indifferent between choosing the optimal value of $t^* = t^O$ and choosing $t^* = 1$ to shut-out the Content Provider.

a) If $\sqrt{\frac{cr^2 + \widehat{C}_1 r + r^2 + 2r + 1}{r^2}} - \frac{1}{r} < b < b^*$ then optimal for network 2 to shut-out the Content Provider, by raising t, eg $t = 1$.

b) If $b^* \le b \le 1 + \widehat{C}_1$ then network 1 announces $t^* = t^O$ given by (S.94).

Note that b^* is very "close"to $1 + \widehat{C}_1$ (informally; i.e., the region is small)

v) These are the only possibilities: if a value of t is chosen so that Case 3 of Theorem III.11 introduces a possible degenerate Nash, network 1 can increase profits by setting $t = 1$.

Subcase: When $b > 1 + \widehat{C}_1$

i) First note that *when a Nash equilibrium exists in Case 3, i.e., in Region 3*, it is always advantageous for network 1 to have the Content Provider use its network. That is, the profit for network 1 is greater with $t = 0$ than shutting the CP out.

ii) Straightforward to show that given $b > 1 + \widehat{C}_1$, $\frac{\partial}{\partial t}\pi_1(p_1^*(t), p_2^*(t), p_3^*(t))|_{t=0} > 0$ and hence if a Nash equilibrium exists with t set, t should be strictly positive.

iii) In this case (i.e., Nash equilibrium and in Region 3) there are two possible cases

a) t can be $t^* = t^O$, i.e., solution to first order conditions

b) $t^* = t^B$ is on the boundary of the Nash equilibrium boundary, the critical point at which network 2 is indifferent between competing with the Content Provider, or lowering its price to drive out the Content Provider.

Sufficient condition for one of these two cases to exist is $1 + c < b \le 2(1+c) + \frac{1}{r}$, (which follows from combining Theorem III.13 with this subcase).

iv) We know from Cor. III.10 than an $\epsilon-$equilibrium exists in this (where $b > 1 + \widehat{C}_1$) in the second stage game if $t = 0$), and hence an "optimal" $\epsilon-$equilibrium also exists -i.e.from above, it follow that $t^* = t^O$ is always an optimal $\epsilon-$equilibrium in Stackelberg game.

v) Under certain conditions, no Nash equilibrium exists in the multistage game *with the Content Provider involved* —i.e., the transfer price is raised to such a high level that it is shut out of the market, (cf Theorems III.9 and III.11 which discuss equilibria for when the Content Provider involved). Cor. III.14 gives a sufficient condition for no Nash equilibrium to exist with $t = 0$, and this implies the only Nash equilibrium is when t is raised to such a level as to shut the Content Provider out of the market.

□

H. Proof of Theorem V.1

THEOREM V.1 For quadratic congestion costs and equal capacities, the optimal strategy for network 1 is to set a negative transfer price $t^* = t^O(b, \widehat{C}_1)$.

Proof. We will show that $\frac{d}{dt}\pi_1^*(t)|_{t=0} < 0$, and that $\pi_1^*(t)$ is concave. When in Case 3 of Lemma III.1, putting $r = 1/b$, gives the defining conditions as

$$\left(\frac{bQ_2}{\widehat{C}_1}\right)^2 - \left(\frac{Q_{11} + bQ_{13}}{\widehat{C}_1}\right)^2 + p_2 - p_1 = p_3 \qquad \text{(S.95)}$$

and $Q_{11} = p_3$, hence Q_{13} is the solution to the quadratic

$$\left(\frac{b(1 - p_3 - Q_{13})}{\widehat{C}_1}\right)^2 - \left(\frac{p_3 + bQ_{13}}{\widehat{C}_1}\right)^2 + p_2 - p_1 = p_3 \quad \text{(S.96)}$$

which reduces to a linear equation, with solution

$$Q_{13} = \frac{b^2(1-p_3)^2 - p_3^2 - \widehat{C}_1^2(p_1 - p_2 + p_3)}{2b(b(1-p_3) + p_3)}. \qquad \text{(S.97)}$$

We have seen that $\pi_1(p_1)$ and $\pi_2(p_2)$ are concave when $r = 1/b$, and that $\pi_3(p_3)$ is in general concave (or quasi-concave), hence we can find the optimum by considering the first order conditions. Equating to zero the derivatives $\frac{\partial_{\pi_i}}{\partial_{p_j}}$ for $i = 1, 2, 3$, and substituting for the partial derivatives by implicitly differentiating (S.96) or (S.97) gives, after simplifying, the equations

$$2b(p_3 + Q_{13}) = \frac{\widehat{C}_1^2(p_1 + t)}{b(1-p_3) + p_3}$$
$$2bQ_2 = \frac{\widehat{C}_1^2 p_2}{b(1-p_3) + p_3}$$
$$(b + t - bt)Q_{13} = \frac{\left(\widehat{C}_1^2 + 2p_3\right)(p_3 - t)}{2b} + b(p_3 - t)(1 - p_3), \qquad \text{(S.98)}$$

whose solution give p_i^*, and Q_i^*. Eliminating p_1 and p_2, and noting that $Q_2 = 1 - p_3 - Q_{13}$, enables us to reduce the set of equations {(S.97), (S.98)} to a pair of simultaneous linear equations in Q_{13}, i.e.,

$$b^2(1-p_3)(3 - 5p_3 - 6Q_{13} - \widehat{C}_1^2(p_3 - t)$$
$$= p_3(b(4p_3 + 6Q_{13} - 2) + p_3)$$

$$\left(\widehat{C}_1^2 + 2p_3\right)(p_3 - t)$$
$$= 2b\left(b\left[(1-t)Q_{13} - (1-p_3)(p_3 - t)\right] + Q_{13}t\right). \qquad \text{(S.99)}$$

Eliminating Q_{13} reduces the equations to a cubic in p_3, where p_3^* is the real root in $[0, 1]$. Explicitly the cubic is

$$6(b-1)^2(b+1){p_3}^3 - (b-1)\left(17b^2 + 7b + 3\widehat{C}_1^2\right){p_3}^2$$
$$+ 2b\left(7b^2 - b + 2\widehat{C}_1^2\right)p_3 - 3b^3 = 0.$$

To show that $\frac{d}{dt}\pi_1^*(t)|_{t=0} < 0$ involves

i) Writing $\pi_1^*(t) = p_1^*(t)\,(p_3^*(t) + Q_{13}^*(t)) + tQ_{13}^*(t)$ and differentiating to obtain $\frac{d}{dt}\pi_1^*(t)|_{t=0}$

ii) Differentiating (S.98) and (S.97) implicitly w.r.t t, and setting $t \mapsto 0$

iii) Showing that when $t = 0$ the resulting set of equations for $\{\pi_1^*(t)|_{t=0}, (p_i^*(0), Q_i^*(0)), (\frac{d}{dt}p_i^*(t)|_{t=0}\,\frac{d}{dt}Q_i^*(t))|_{t=0})\}$ do not have a consistent solution unless $\frac{d}{dt}\pi_1^*(t)|_{t=0} < 0$. This is proved using Mathematica [18].

□