

# Network Pricing and Rate Allocation with Content Provider Participation

Prashanth Hande<sup>1,2</sup>, Mung Chiang<sup>1</sup>, Robert Calderbank<sup>1</sup>, Sundeep Rangan<sup>2</sup>

<sup>1</sup>Department of Electrical Engineering, Princeton University, NJ 08544, USA

<sup>2</sup>Qualcomm Flarion Technologies, NJ 08807, USA

**Abstract**—Pricing content-providers for connectivity to end-users and setting connection parameters based on the price is an evolving model on the Internet. The implications are heavily debated in telecom policy circles, and some advocates of “Network Neutrality” have opposed price based differentiation in connectivity. However, pricing content providers can possibly subsidize the end-user’s cost of connectivity, and the consequent increase in end-user demand can benefit ISPs and content providers. This paper provides a framework to quantify the precise trade-off in the distribution of benefits among ISPs, content-providers, and end-users.

The framework generalizes the well-known utility maximization based rate allocation model, which has been extensively studied as an interplay between the ISP and the end-users, to incorporate pricing of content-providers. We derive the resulting equilibrium prices and data rates in two different ISP market conditions: competition and monopoly. Network neutrality based restriction on content-provider pricing is then modeled as a constraint on the maximum price that can be charged to content-providers. We demonstrate that, in addition to gains in total and end-user surplus, content-provider experiences a net surplus from participation in rate allocation under low cost of connectivity. The surplus gains are, however, limited under monopoly conditions in comparison to competition in the ISP market.

**Keywords:** Pricing, Network Neutrality, ISP Market, Rate Allocation, Optimization, Distributed Algorithm.

## I. INTRODUCTION

Internet Service Providers (ISP) provide a network over which end-users and content (or application) providers connect for data exchange. The connectivity between the end-users and the content-providers is characterized by a set of parameters, a predominant component of which is data rate allocation.

The increase in Internet activity is placing further demand on network capacity for improved connectivity between end-users and content-providers. On wireline networks, ISPs have voiced their intention to charge content-providers to recover the cost of investment on network capacity, and provide differential allocation of connectivity parameters based on the payment received. Some proponents of “Network Neutrality” [1]–[4] are opposed to the idea and claim that content-providers are adversely impacted from such differential pricing, reducing competition in the content market. On the wireless side, access networks traditionally lacked sufficient data capacity, and wireless ISPs closely controlled the content delivered on the network. However, with increasing data capacity on new generation wireless networks, end-users are increasingly demanding content and applications similar to those on wireline networks. The wireless

ISPs are forced to relinquish tight control of content and instead look into business models where content-providers can be charged for provisioning of connectivity parameters [5].

The emerging Internet pricing models, where the price of connectivity is shared between content-providers and end-users, offer economic surplus trade-offs that remain to be quantified. Intuitively, the end-users pay less for connectivity and the ISPs have an additional source of revenue. More surprisingly, the content-providers may also stand to gain from such pricing models if subsidizing connectivity to end-users translates into revenue gains in the content market. However, the content-providers are unsure of the extent to which their payment is passed on as a subsidy to end users. In fact, the precise gains to the three entities will depend upon the complex interplay among them. Our work begins a systematic study on this topic by analyzing rate allocation resulting from this three-way interaction.

The following are two sample scenarios that can be analyzed through the model developed here:

- A wireless ISP is planning to allow high-resolution video content hosted by a third-party content-provider on its network. The ISP is unsure if the users are willing to pay for the entire cost of delivering the video. Instead, the wireless ISP is deciding on a pricing scheme where the content-provider shares the price of video delivery with the users. The content-provider can gain from inserting advertisements in the video but is unsure about how much it can pay the ISP and remain profitable.
- A wireline ISP is planning on upgrading to a fiber based access network. To recover the cost, the ISP plans to charge both the end-users and the content-providers, and needs to determine the appropriate price split. Content-providers have approached the regulators to prevent the ISP from charging them. The regulatory authority will have to make a ruling on the extent to which the ISP can charge the content-provider.

Rate allocation has been extensively studied in other contexts before. The work in [6] proposed a model, the basic network utility maximization model (NUM), for network rate allocation. Based on this model, the work in [7] proposed a mechanism (referred to as *dual-gradient mechanism*) to achieve the equilibrium rates. The mechanism determines the price charged by the ISP (*price level*) in a competitive ISP market, resulting in maximization of the sum of end-user utilities. However, the model cannot determine the precise split of the price charged to the end-user and the content-provider (*price structure*), when

the content-provider shares with the end-user the utility derived from a traffic flow. In addition, the model is not applicable to monopoly ISP market where the price setting maximizes the ISP profit. We extend the NUM model to include content-provider pricing, and establish the equilibrium price structure and rate allocation, in both competitive and monopoly ISP markets. We also propose a mechanism that augments the dual-gradient mechanism and converges to the equilibrium price structure in a competitive ISP market.

There are many aspects to the “Network Neutrality” debate beyond the scope of this paper’s focus, but a central issue in the debate is indeed the pricing of content-provider and the consequent surplus distribution between the ISP, the end-users, and the content-providers. Content and connectivity are two different commodities under transaction in this interaction, with neither of them a meaningful product without the other. It is difficult to measure the end-user demand for content and connectivity separately. Increasingly though, content is being made available free of charge to end-users, with content-providers earning revenue through alternate strategies, such as advertising. This allows us to model the end-user’s joint demand for content and connectivity, as a function of the connectivity price. In addition, the content market has a large number of content-providers each with a relatively small market share, displaying a “long-tail” distribution. This allows us to model the content-provider as a price-taker with a demand on connectivity, as a function of the connectivity price.

A number of recent works in economics [8], [9] have proposed quantitative models on network neutrality but do not deal with the engineering aspect of rate allocation or the specifics of connectivity market. Instead, they focus on externality effects and the “two-sided model” [10]. The two-sided model considers the content-provider and the end-user as two sides that interact in a market enabled by the platform (network) provider, the ISP. The total volume of the market is influenced by the price structure rather than the price level alone. The models in [8]–[10] assume a utility for each side that is a function of the number of members on the other side (the network effect). The appropriate price split between the two sides under different platform market conditions is then analyzed. Our model is different from that considered in [10] in that the utility here is a function of the data rate. The analysis is still “two-sided” since the price structure, in addition to the price level, determines the rate allocation. This model is more relevant in the ISP market where end-users typically access data from a small number of content-providers and derive no direct benefit from the overall number of content-providers. Rather, the end-users benefit from the extent of interaction with each content-provider, as measured by the flow data rate. Our analysis reveals the pricing structure for rate allocation under ISP competition and monopoly, and provides a quantitative framework for this aspect of the network-neutrality debate.

The paper is organized as follows. Content provider is introduced into the rate allocation model in Section II. The equilibrium data rates and price structures are established for a competitive ISP market and a mechanism to attain the equilibrium is proposed in Section III. The equilibrium for a monopoly ISP market is established in Section IV. In section V, we

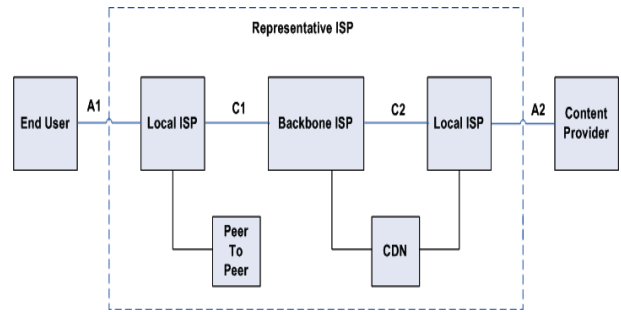


Fig. 1. ISP, End User, Content Provider Interaction

discuss practical considerations and illustrate the implication of content-provider participation through a simple toy-example. Network neutrality related restriction and consequent impact are analyzed in Section VI before we conclude in Section VII.

## II. RATE ALLOCATION MODEL

Consider a network of data links over which content-providers (CP) and end-users (EU) exchange data. Let  $f(i, j)$  denote data flow between CP  $j$  and EU  $i$  and let  $x_f$  be the associated flow data rate. Each flow corresponds to the interaction between a CP-EU pair. We denote by  $U(i)$  the set of flows serving EU  $i$  and by  $C(j)$  the set of flows serving CP  $j$ . Let  $u_f(x_f)$  and  $v_f(x_f)$  be the utilities associated with flow  $f(i, j)$  for EU  $i$  and CP  $j$ , respectively. We make the following standard assumption on the utility functions:

*Assumption 1:* The utility functions  $u_f(\cdot)$  and  $v_f(\cdot)$  are increasing, strictly concave, and twice continuously differentiable.

Multiple ISPs, interacting through peering and transit agreements, can be operating the links between the CP and the EU. We consider a representative ISP pricing the EU and the CP as shown in Figure 1. The representative ISP’s revenue can then be viewed as split among these multiple entities based on models [11] representing the interactions between them.

Let  $p_f(i, j)$  and  $q_f(i, j)$  be the prices per unit data rate on flow  $f$  charged by the ISP to EU  $i$  and CP  $j$ , respectively. Let  $L(f)$  denote the set of links used by flow  $f$ . Capacity of link  $l$  is  $c_l$  and the cost to provision link capacity is  $w_l$  per unit capacity. If capacity cost function is strictly convex,  $w_l$  can be replaced by the marginal cost of capacity. In the special case of fixed capacity,  $w_l$  appears in distributed algorithms for rate allocation as a set of prices that have the interpretation of Lagrangian dual variables [6]. A major contribution of the dual-gradient mechanism was the ability to determine the appropriate prices in the fixed capacity case, without monitoring individual flows on the links. Fixing  $w_l$ , as opposed to fixing capacity, allows us to focus on the price structure, the price split between the EU and the CP. We present our analysis for the case of fixed  $w_l$ . However, the analysis in the rest of this paper determines the price structure even with fixed capacities, with the price level determined through the dual-gradient mechanism.

The ISP, EUs, and CPs interact through prices  $\{p_f\}$  and  $\{q_f\}$  to decide equilibrium data rates  $\{x_f\}$  that can be supported in the ISP network with an appropriate capacity provisioning  $\{c_l\}$ . Denote by  $y_f(p_f)$  the data rate on flow  $f(i, j)$  demanded by

EU  $i$  in response to price  $p_f$  to maximize net utility

$$\text{EU: } \begin{array}{ll} \text{maximize} & \sum_{f \in U(i)} (u_f(x_f) - p_f x_f) \\ \text{variables} & \{x_f \geq 0\} \end{array} \quad (1)$$

This implies that

$$y_f(p_f) = u_f'^{-1}(p_f) \quad (2)$$

Correspondingly, CP  $j$  demands data rate  $z_f(q_f)$  on flow  $f(i, j)$  that maximizes net revenue

$$\text{CP: } \begin{array}{ll} \text{maximize} & \sum_{f \in C(j)} (v_f(x_f) - q_f x_f) \\ \text{variables} & \{x_f \geq 0\} \end{array} \quad (3)$$

resulting in a demand function

$$z_f(q_f) = v_f'^{-1}(q_f) \quad (4)$$

Let  $\mu_f = \sum_{l \in L(f)} w_l$  be the aggregate cost of provisioning capacity along the route corresponding to flow  $f$ . We assume that the demand at zero price on either side (CP or EU) exceeds the demand at the capacity cost on the other side.

*Assumption 2:* Demand functions  $\{y_f(p_f), z_f(q_f)\}$  are such that  $z_f(\mu_f) < y_f(0)$  and  $z_f(0) > y_f(\mu_f)$ .

The end-user elasticity of demand  $\eta_f^E$  and the content-provider elasticity of demand  $\eta_f^C$  represent the price sensitivity of the demand functions. They are defined as:

$$\begin{aligned} \eta_f^E &= -\frac{y_f'(p_f)p_f}{y_f(p_f)} \\ \eta_f^C &= -\frac{z_f'(q_f)q_f}{z_f(q_f)} \end{aligned} \quad (5)$$

We now analyze the interaction between the three entities and establish the resulting equilibrium data rates and prices, both under a competitive ISP market and a monopoly ISP market.

### III. COMPETITIVE ISP MARKET

In a competitive ISP market, the ISPs have no control on the prices but instead accept prices that are market driven. With end-users and content-providers modeled as price takers, the prices  $\{p_f, q_f\}$  are exogenous to all the three entities involved.

#### A. Equilibrium Prices and Data Rates

An ISP in a competitive market supports data rates  $x_{f(i,j)}$  and provisions capacities that maximize the profit at the given market price:

$$\text{ISP: } \begin{array}{ll} \text{maximize} & \sum_f (p_f + q_f)(x_f) - \sum_l w_l c_l \\ \text{subject to} & \sum_{f: l \in L(f)} x_f \leq c_l, \quad \forall l \\ \text{variables} & \{x_f \geq 0, c_l \geq 0\} \end{array} \quad (6)$$

If  $r_f = p_f + q_f$  is the price level per unit flow charged by the ISP, it can be shown that the ISP maximization problem (6) can be simplified to

$$\begin{array}{ll} \text{maximize} & \sum_f (r_f - \mu_f)(x_f) \\ \text{variables} & x_f \geq 0 \end{array} \quad (7)$$

with the ISPs provisioning just enough capacity to support the maximizing data rate. The maximization problem (7) has a linear objective and has a simple solution with an intuitive interpretation. If  $r_f < \mu_f$ , then the ISPs will not accept any non-zero data rate and the market will not sustain in this case. If  $r_f > \mu_f$ , then the ISPs maximize the accepted data rate. If

$r_f = \mu_f$ , then the ISP maximization problem accepts any data rate as a solution. This creates the possibility of a continuum of equilibrium points, common in maximization of a linear objective. To get around the difficulty, we make the following reasonable assumption:

*Assumption 3:* If  $r_f = \mu_f$  resulting in zero profit at all acceptable data rates, the ISPs accept data rates that maximize revenue.

At equilibrium, the acceptable data rate to the ISP should equal the net demand by the EU and the CP. The net data rate demand by the EU and the CP at any price structure  $\{p_f, q_f\}$  is simply the minimum of the two demand functions since that is the maximum rate that is acceptable to both the EU and the CP. Given  $\{y_f(p_f), z_f(q_f)\}$ , this implies that the ISP faces a joint demand of  $\min\{y_f(p_f), z_f(q_f)\}$ . Assumption 3 implies that ISPs maximize the acceptable data rates at prices  $r_f \geq \mu_f$ . For a given price level  $r_f$ , the market price structure  $\{p_f, q_f\}$  is determined by the following maximization problem, with optimal value denoted as  $x_f(r_f)$ :

$$\begin{array}{ll} \text{maximize} & \min\{y_f(p_f), z_f(q_f)\} \\ \text{subject to} & p_f + q_f = r_f \\ \text{variables} & p_f \geq 0, q_f \geq 0 \end{array} \quad (8)$$

From the perspective of the system consisting of EUs, ISPs and CPs, the rate allocation needs to satisfy the maximization of the total system surplus, and is given by

$$\begin{array}{ll} \text{SYSTEM: } \text{maximize} & \sum_f (u_f(x_f) + v_f(x_f)) - \sum_l w_l c_l \\ \text{subject to} & \sum_{f: l \in L(f)} x_f \leq c_l, \quad \forall l \\ \text{variables} & \{x_f \geq 0, c_l \geq 0\} \end{array} \quad (9)$$

The set of optimization problems considered here is similar to the ones considered in [6] but now augmented by the CP maximization problem. It can be seen that the three set of problems (1), (3), and (6) admit any solution where the price level is  $r_f \geq \mu_f$  and the price structure  $\{p_f, q_f\}$  is determined by the maximization problem (8). However, any ISP charging  $r_f > \mu_f$  is undercut by a competing ISP, resulting in the price level  $r_f = \mu_f$  being the only sustainable equilibrium. The following theorem proved in Appendix A characterizes the price structure at this equilibrium.

**Theorem 1:** The equilibrium prices  $\{\hat{p}_f, \hat{q}_f\}$  and the corresponding equilibrium data rates  $\hat{x}_f$  that simultaneously solve (1), (3), and (6) for  $r_f = \mu_f$  is given by:

$$\begin{aligned} \hat{p}_f + \hat{q}_f &= \mu_f \\ y_f(\hat{p}_f) &= z_f(\hat{q}_f) \end{aligned} \quad (10)$$

The equilibrium data rate demand faced by the ISP is  $\hat{x}_f = y_f(\hat{p}_f) = z_f(\hat{q}_f)$  and the capacity provisioned will be the minimum required to support the equilibrium data rate:  $\hat{c}_l = \sum_{f: l \in L(f)} \hat{x}_f$ . Further, at the equilibrium with  $r_f = \mu_f$ , (9) is also solved so that the system surplus is also maximized.

The theorem is similar to that in [6] and is related to the extensively studied first welfare theorem and the existence theorem of general equilibrium in economics [19]. The theorem is stated for a price level  $r_f = \mu_f$ , which we argued as the only stable equilibrium, but the price structure at any price level  $r_f \geq \mu_f$  can be obtained along similar lines. Fixing

$r_f = \mu_f$  allows us to focus on the price structure and the demand functions rather than the price level and the utility functions as in the existing NUM literature.

### B. Mechanism to Attain Equilibrium

In the absence of CP, the ISP can indicate price  $r_f = \mu_f = \sum_{l \in L(f)} w_l$ , resulting in EU demand  $x_f = y_f(r_f)$ . Any deviation from this mechanism by the ISP results in EUs switching to other more competitive ISPs, and any deviation from the mechanism by the EU results in loss of net utility to the EU. The dual-gradient mechanism is a variation of the above mechanism [7], [15], [17] for the case where capacity is fixed and the cost  $w_l$  is adjusted to suit the fixed capacity available. In particular, the following time update of the link cost is known to converge to the appropriate price level:

$$\begin{aligned} w_l[t+1] &= \left[ w_l[t] - \alpha[t] \left( c_l - \sum_{f: l \in L(f)} x_f(r_f[t]) \right) \right]^+, \quad \forall l \\ r_f[t] &= \sum_{l \in L(f)} w_l[t] \end{aligned} \quad (11)$$

Here  $\alpha[t]$  is an appropriate step size diminishing in time. This pricing update is a single price update per link based on *aggregate* demand at the link as opposed to a scheme requiring price update for individual flows.

The primary difficulty in extending the above algorithm to include CP participation in rate allocation is to determine the appropriate split of the ISP demanded price  $r_f$  (or  $r_f[t]$  if capacity is fixed). The split between the CP and the EU needs to be determined without the ISP having to monitor the individual EU and CP utilities and demand functions. Let  $\theta_f$  denote the fraction of the total price borne by the EU so that  $p_f = \theta_f r_f$ . Let  $\alpha[t] = 1/t$ . We propose the following algorithm to achieve the equilibrium under CP participation:

#### Algorithm 1 (Rate Allocation Mechanism with CP):

Initialize:  $\theta_f[0] = 0$

- 1) ISP indicates  $r_f[t]$  as the price per unit data rate to the EU.
- 2) EU picks up  $p_f[t] = \theta_f[t] r_f$  of the total price with

$$\theta_f[t+1] = B(\theta_f[t] + \alpha[t] (y_f(p_f[t]) - z_f(q_f[t]))) \quad (12)$$

where  $B(\theta) = \max(\min(\theta, 1), 0)$  is a bounding function.

- 3) CP picks up  $q_f[t] = r_f - p_f[t]$ .
- 4) EU demands data rate  $y_f(p_f[t])$ , CP demands data rate  $z_f(q_f[t])$ .
- 5) ISP calculates net demand as  $x_f[t] = \min\{y_f(p_f[t]), z_f(q_f[t])\}$ .
- 6) If capacity cost is fixed,  $r_f[t] = \mu_f$ . If capacity is fixed,  $r_f[t]$  is updated according to (11).

Continue:  $t := t + 1$

The novelty in the algorithm is the fraction update as given in (12) that achieves the equilibrium split of ISP indicated price between the EU and the CP. The following theorem showing convergence of the algorithm is proved in Appendix B.

**Theorem 2:** Algorithm 1 converges to the equilibrium price structure as given in Theorem 1.

The algorithm requires the end-user to monitor the per-flow content-provider demand at any given time  $t$ . The ISP update of the link capacity (or link price in the case of fixed capacity) is based on aggregate demand as in the basic dual-gradient mechanism. One possible implementation of the algorithm in a packet flow from the content-provider to the end-user is for the links to indicate their prices in the data packets of the packet stream and for the end-user to indicate the content-provider's fraction of the total route price in the ACK packets.

## IV. MONOPOLY ISP MARKET

In a monopoly market, the prices are under the complete control of the monopolist ISP and it is assumed that the ISP knows the demand functions of the end-users and the content-providers. This means that the ISP determines the flow data rate  $x_f$  and the prices  $\{p_f, q_f\}$  at which the ISP profit is maximized. For a given price  $r_f$ , the ISP has to split the price appropriately to achieve the maximum possible demand  $x_f(r_f)$  as given in (8). Further, the total price  $r_f$  has to be set so as to maximize the ISP profit, resulting in the following maximization problem:

$$\text{MONOPOLY ISP: } \begin{aligned} &\text{maximize} && \sum_f (r_f - \mu_f)(x_f(r_f)) \\ &\text{variables} && r_f \geq 0 \end{aligned} \quad (13)$$

The ISP price level is determined by the first-order condition that has the well known form:

$$\frac{r_f - \mu_f}{r_f} = 1/\eta_f \quad (14)$$

where  $\eta_f = -\frac{x'_f(r_f)r_f}{x_f(r_f)}$  is the elasticity of the combined demand faced by the monopolist ISP. The monopoly ISP price structure  $\{\tilde{p}_f, \tilde{q}_f\}$  is determined by the following theorem, proved in Appendix C.

**Theorem 3:** The set of prices  $\{\tilde{p}_f, \tilde{q}_f, \tilde{r}_f\}$  and data rates  $\{\tilde{x}_f\}$  that solve (13) satisfy:

$$\begin{aligned} \tilde{p}_f(1 - 1/\eta_f^E) + \tilde{q}_f(1 - 1/\eta_f^C) &= \mu_f \\ \tilde{x}_f = y_f(\tilde{p}_f) = z_f(\tilde{q}_f) \end{aligned} \quad (15)$$

The capacity provisioned will be the minimum required to support the equilibrium data rate:  $\tilde{c}_l = \sum_{f: l \in L(f)} \tilde{x}_f$ .

The theorem is a statement on the price structure for a monopoly ISP. The monopoly price level  $\tilde{r}_f = \tilde{p}_f + \tilde{q}_f$  satisfies  $\tilde{r}_f = \mu_f/(1 - 1/\eta)$ , and is well-known [19] to have the interpretation that a demand elasticity larger than 1 results in finite monopoly prices, whereas a demand elasticity less than or equal to 1 results in infinite monopoly prices and near-zero demand. The derived price structure has a similar interpretation with higher elasticities of demand resulting in lower prices.

If the monopoly ISP has incomplete knowledge of the end-user and content-provider demand functions, then mechanisms to achieve the monopoly price structure need to be designed. While we have presented the equilibrium price structure and data rates, the trade-off in benefits between the end user, content-provider, and the ISP remains to be explored. We will quantify this trade-off in Section VI and discuss the implication of content-provider participation in rate allocation.

## V. PRACTICAL CONSIDERATIONS AND ILLUSTRATION

We derived results for two extreme ISP market conditions, monopoly and perfect competition. In practice, ISP markets conditions in many parts of the world are in between the two extremes. The large capital expenditure required to setup an ISP typically results in oligopolistic market structures, where a few firms dominate the market. In general, oligopolistic markets have equilibrium prices and usage in between that of competition and monopoly, and the results in this paper serve as a suitable guideline to analyze the resulting market equilibrium. Under certain models, however, oligopoly markets result in non-intuitive equilibrium [12], in which case such markets will have to be explicitly analyzed for additional insights.

We also assume a flow dependent, route dependent and usage dependent pricing on the demanded data rates. In practice, pricing is rarely flow or route dependent, and is sometimes based on total volume of data rather than data rates. Flat-rate pricing models, where prices stay flat up to a threshold data consumption, are not uncommon either, especially in the U.S. wireline market today. Several papers have addressed the comparison with usage dependent pricing [13], [14]. Translating prices on data rates to that on volume of data can be accommodated by considering the time period over which the volume of data is calculated, and a multiplexing factor that captures the effect of the volume of data on capacity requirements. Imposing flow and route independence on the traffic is harder to accomplish, and will be the target of future work.

We now provide a simple example to illustrate the dependency of the equilibrium values on the content provider utility functions, under ISP competition and monopoly. The illustration brings out two salient points related to the rate allocation model. Content provider participation can significantly increase the data rates under both ISP monopoly and competition. Monopoly ISP market conditions dramatically reduce the equilibrium data rates in comparison to a competitive ISP market.

Consider EU utilities  $u_f(x_f)$ , parameterized by  $\{a_f, \alpha_f\}$ , of the form

$$u_f(x_f) = \begin{cases} a_f \log(x_f) & \text{if } \alpha_f = 1, \\ a_f(1 - \alpha_f)^{-1} x_f^{1-\alpha_f} & \text{if } 0 \leq \alpha_f < 1 \end{cases} \quad (16)$$

and CP utilities  $v_f(x_f)$ , parameterized by  $\{b_f, \beta_f\}$ , given by

$$v_f(x_f) = \begin{cases} b_f \log(x_f) & \text{if } \beta_f = 1, \\ b_f(1 - \beta_f)^{-1} x_f^{1-\beta_f} & \text{if } 0 \leq \beta_f < 1 \end{cases} \quad (17)$$

The utility functions result in end-user demand function  $y_f(p_f)$  and content-provider demand function  $z_f(q_f)$  given by

$$y_f(p_f) = (p_f/a_f)^{-1/\alpha_f}, \quad z_f(q_f) = (q_f/b_f)^{-1/\beta_f} \quad (18)$$

The elasticities of the demand functions are given by

$$\eta^E = 1/\alpha_f, \quad \eta^C = 1/\beta_f \quad (19)$$

The assumed utility functions for the EU and CP results in constant elasticity of demand.

We look into a single flow for this illustration and use the following nominal parameters  $\{a_f = \$50, \alpha_f = 0.5, b_f =$

TABLE I  
BROADBAND PRICING AND USAGE VS. UTILITY LEVEL  
 $a_f = 50, \alpha_f = 0.5, \beta_f = 0.5$

$b_f$	ISP Competition		ISP Monopoly	
	EU Price	EU Demand	EU Price	EU Demand
0	\$50	1.0 Mbps	\$100	250 kbps
20	\$36	2.0 Mbps	\$71	490 kbps
40	\$28	3.25 Mbps	\$56	810 kbps
60	\$23	4.84 Mbps	\$45	1.21 Mbps
80	\$19	6.76 Mbps	\$38	1.69 Mbps
100	\$17	9.0 Mbps	\$33	2.25 Mbps

TABLE II  
BROADBAND PRICING AND USAGE VS. UTILITY CURVATURE  
 $a_f = 50, b_f = 50, \alpha_f = 0.5$

$\beta_f$	ISP Competition		ISP Monopoly	
	EU Price	EU Demand	EU Price	EU Demand
0.2	\$6	63 Mbps	\$33	2.35 Mbps
0.4	\$14	13 Mbps	\$45	1.24 Mbps
0.6	\$19	7 Mbps	\$55	824 kbps
0.8	\$22	5 Mbps	\$68	547 kbps
1.0	\$25	4 Mbps	\$100	250 kbps
1.05	\$26	3.8 Mbps	\$151	110 kbps

$\$50, \beta_f = 0.5\}$ . The aggregate capacity cost per Mbps of data rate is assumed to be  $\mu_f = \$50$  per month. Table I shows the equilibrium data rate and the price charged to the end-user at different content-provider utility levels  $b_f$ . CP does not participate in rate allocation when  $b_f = 0$ . Data rates increase from 1Mbps to 9Mbps as  $b_f$  increases to \$100 under ISP competition. The end-user price decreases to \$17 per month as opposed to the \$50 per month paid without content provider contribution. For the monopoly case, the increase in broadband data rates and decrease in end-user price is less dramatic, rates increase to 2Mbps, and end-user price decreases to \$33 for  $b_f = \$100$ .

Table II shows the equilibrium data rate, and the end-user price against  $\beta_f$ , the inverse of the CP elasticity of demand. Higher the  $\beta_f$ , lower the demand elasticity, indicating less willingness to pay. This results in high end-user prices and low

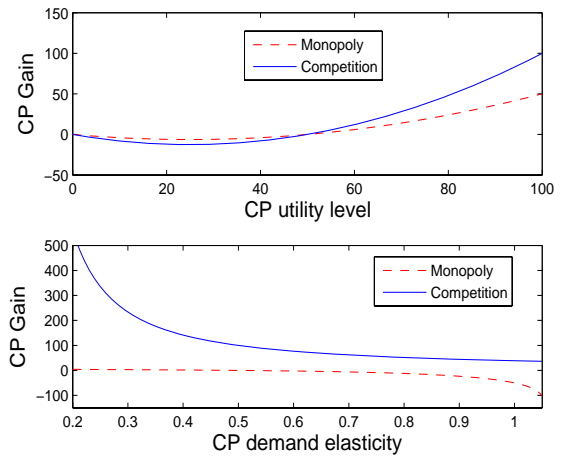


Fig. 2. Plot of content-provider gain from participation in rate allocation vs.  $b_f$  and  $\beta_f$ . Content-provider realizes positive gains for high  $b_f$  and low  $\beta_f$  under both ISP competition and monopoly. Gains are limited under monopoly.

broadband data rates. In fact, with very low content-provider elasticity of demand ( $\beta_f \geq 1$ ), the end-user price increases and the data rate decreases, beyond the equilibrium values without CP participation. The ISP captures all of the surplus resulting from content-provider participation, resulting in adverse impact on data rates and end-user prices.

Figure 2 plots the content-provider gain in net-utility from participation in rate allocation. The gain is obtained by calculating the difference between the net utility under the equilibrium with content-provider payment, and the net-utility under the equilibrium with only the end-user paying for connectivity. It can be observed that high content-provider utility levels ( $b_f$ ) result in significant gain for the content-provider from participation in rate allocation. Low content provider demand elasticity (high  $\beta_f$ ), on the other hand, results in decreasing gains for the content-provider. The gains are comparatively limited under monopoly, with the ISP capturing much of the surplus resulting from content-provider participation.

The illustration confirms the intuition that the end-user stands to gain from content-provider participation. The precise gains are quantified making it apparent that the gains are limited under ISP monopoly conditions. It is less intuitive that the content-provider can gain from participation in rate allocation, as demonstrated in the above example. We further explore the distribution of surplus across the end-user, content-provider, and the ISP in the next section.

## VI. IMPLICATIONS ON NETWORK NEUTRALITY

“Network Neutrality” apparently is a term with no clear definition and too many intertwined, contentious aspects. Interpretations of the term ranges from restrictions on traffic prioritization and content-provider pricing to setting boundaries of vertical integration in the content and connectivity markets. We consider the content-provider pricing aspect of “network neutrality” and model the related restriction as a constraint on the maximum price  $\bar{q}$  that the ISP can charge the content-provider. Policy makers are faced with questions on the trade-off in benefits from imposing such restrictions. Relevant policy questions include the following:

- What is the gain in surplus to the end-users due to content-provider participation?
- Does the ISP reap the surplus at the expense of the content-provider?
- Does the content-provider always experience a loss in surplus? Is it possible that the content provider participation can induce increased usage from end-user, sufficient enough to result in a net surplus increase to the content-provider?
- Is there a gain in total surplus for the whole system?

We can quantify the trade-off in the surpluses through the rate allocation framework presented in sections III and IV. This does not prescribe, nor mandate, any particular position along the wide spectrum of conclusions on content provider pricing. What it does is to provide an unambiguous language for a more meaningful dialog on this aspect of the net-neutrality debate and to allow a policy maker to make decisions based on well-quantified results. We discuss below implications when the constraint price varies between zero and the equilibrium

content-provider price. We suppress the subscript  $f$  in this section for notational simplicity.

### A. Net Neutrality in a Competitive ISP Market

Consider a competitive ISP market with a content-provider price restriction below the market equilibrium:  $0 \leq \bar{q} \leq \hat{q}$ . Since the end-user and content-provider demand are equal at the competitive market equilibrium, the CP price restriction below  $\hat{q}$  results in a larger share of the price to the end-user. The end-user demand is consequently less than the possible content-provider demand at price  $\bar{q}$ , and will be the prevailing data rate demand at CP price restriction  $\bar{q}$ . The end-user surplus  $S_C^E(\bar{q})$ , content-provider surplus  $S_C^C(\bar{q})$ , the ISP surplus  $S_C^I(\bar{q})$ , and total surplus  $S_C^T(\bar{q})$  are then given by

$$\begin{aligned} S_C^E(\bar{q}) &= u(y(\mu - \bar{q})) - (\mu - \bar{q})y(r - \bar{q}) \\ S_C^C(\bar{q}) &= v(y(\mu - \bar{q})) - \bar{q}y(\mu - \bar{q}) \\ S_C^I(\bar{q}) &= 0 \\ S_C^T(\bar{q}) &= u(y(\mu - \bar{q})) + v(y(\mu - \bar{q})) - \mu y(\mu - \bar{q}) \end{aligned} \quad (20)$$

A “net-neutrality” policy can be parameterized by the restrictions imposed on the CP prices  $\bar{q}$ . Different “degree of neutrality” is reflected in the strictness of such restrictions. A number of papers including [8], [9] model the CP price as equal to zero ( $\bar{q} = 0$ ) in the net-neutral regime, which is not necessarily representative of the present-day situation since content-providers pay non-zero network access charges to the ISPs. A more precise analysis is to look into the changes in the surpluses as  $\bar{q}$  varies from 0 to  $\hat{q}$ . The following theorem establishes the first-order derivatives that provides insight into this variation.

**Theorem 4:** The partial derivatives of the surpluses  $S_C^E(\bar{q})$ ,  $S_C^C(\bar{q})$ ,  $S_C^T(\bar{q})$  with respect to  $\bar{q}$ , for  $0 \leq \bar{q} \leq \hat{q}$ , is given by:

$$\begin{aligned} \frac{\partial S_C^E(\bar{q})}{\partial \bar{q}} &= y \\ \frac{\partial S_C^C(\bar{q})}{\partial \bar{q}} &= \frac{y}{\mu - \bar{q}} ((v'(y) - \bar{q})\eta^E - (\mu - \bar{q})) \\ \frac{\partial S_C^T(\bar{q})}{\partial \bar{q}} &= \frac{y\eta^E}{\mu - \bar{q}} (v'(y) - \bar{q}) \end{aligned} \quad (21)$$

where  $y = y(\mu - \bar{q})$  is the equilibrium end-user demand under the price restriction  $\bar{q}$ . The ISP surplus  $S_C^I(\bar{q})$  is uniformly zero.

It is apparent from (21) that the end-user surplus increases when the CP price restriction is relaxed. The ISP surplus remains zero indicating complete transfer of surplus from the content-provider to the end-user. The term  $v'(y) - \bar{q} \geq 0$  for  $\bar{q} \leq \hat{q}$  indicating that the total surplus increases, which is no surprise since a competitive market maximizes the total surplus and any price restriction can only worsen the total surplus.

The non-intuitive result, however, is the behavior of content-provider surplus. It is seen to increase with relaxation of the price restriction, if the end-user price sensitivity  $\eta^E$  is sufficiently high, or the cost of provisioning capacity,  $\mu$ , is sufficiently low. End users with high price sensitivity consume sufficiently more data rate at prices subsidized by the content-provider that the subsidizing content-provider also sees an increase in surplus. Lower the cost of provisioning capacity,

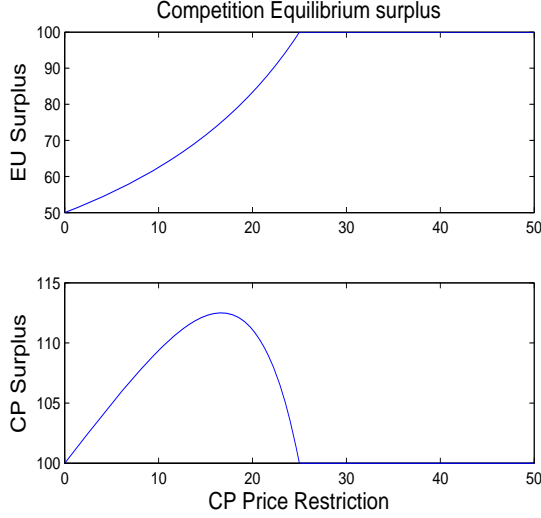


Fig. 3. Competitive Equilibrium end-user and content-provider surplus vs.  $\bar{q}$ . End user surplus increases with  $\bar{q}$  and content-provider surplus peaks at a value of  $\bar{q}$  less than the equilibrium value  $\hat{q}$ .

lower is the end-user price sensitivity required for the content provider to benefit from participation in rate allocation. The content-provider surplus can be seen to be maximized at a value of  $\bar{q}$  less than the equilibrium price  $\hat{q}$ . Figure 3 is a plot of the end-user and content-provider surplus against  $\bar{q}$  for the example illustrated in Section V with  $\{a_f = \$50, \alpha_f = 0.5, b_f = \$50, \beta_f = 0.5\}$ .

In summary, in a competitive ISP market:

- The participation of the content-provider in rate allocation increases the end-user surplus, and the total surplus, while the ISP surplus remains unchanged.
- The content-provider can increase surplus through participation in rate allocation if the end-user price sensitivity is sufficiently high relative to the cost of connectivity.
- Lower cost of connectivity reduces the end-user price sensitivity required for the content-provider to gain from participation in rate allocation.

### B. Net Neutrality in a Monopoly ISP Market

Consider now the monopoly ISP case. The end-user surplus  $S_M^E(\bar{q})$ , content-provider surplus  $S_M^C(\bar{q})$ , the ISP surplus  $S_M^I(\bar{q})$ , and total surplus  $S_M^T(\bar{q})$  are given by

$$\begin{aligned} S_M^E(\bar{q}) &= u(y(r - \bar{q})) - (r - \bar{q})y(r - \bar{q}) \\ S_M^C(\bar{q}) &= v(y(r - \bar{q})) - \bar{q}y(r - \bar{q}) \\ S_M^I(\bar{q}) &= (r - \mu)y(r - \bar{q}) \\ S_M^T(\bar{q}) &= u(y(r - \bar{q})) + v(y(r - \bar{q})) - \mu y(r - \bar{q}) \end{aligned} \quad (22)$$

where  $r = r(\bar{q})$  is the price charged by the monopoly ISP as a function of the content-provider price restriction.

To quantify the surplus behavior, we need to evaluate  $r(\bar{q})$ . To simplify the exposition, we assume that the end-user and content-provider utility functions are such that the elasticities of demand  $\eta^E$  and  $\eta^C$  are fixed and greater than 1. For low values of  $\bar{q}$ , it is apparent that the content-provider's demand  $z(\bar{q})$  exceeds the end-user demand  $y(r - \bar{q})$ . The monopoly ISP

determines the price  $r(\bar{q})$  through the following maximization

$$\begin{aligned} &\text{maximize} && \sum_f (r - \mu)(y(r - \bar{q})) \\ &\text{variables} && r \geq 0 \end{aligned} \quad (23)$$

resulting in  $r(\bar{q}) = \frac{\eta^E \mu - 1}{\eta^E - 1}$ . As  $\bar{q}$  increases, the content-provider demand falls until it equals the end-user demand generated by the monopoly determined price  $r(\bar{q})$ . This threshold price,  $q^t$ , at which the two demands are equal, can be determined by solving  $y(r(q^t) - q^t) = z(q^t)$ . It can be verified that  $q^t \leq \tilde{q}$ , where  $\tilde{q}$  is the equilibrium content-provider price under no restriction. For  $\bar{q} > q^t$ , the monopoly ISP can maximize revenue by equating the two demands so that  $r(\bar{q}) = y^{-1}(z(\bar{q})) + \bar{q}$ . We have

$$r(\bar{q}) = \begin{cases} \frac{\eta^E \mu - 1}{\eta^E - 1} & \text{if } \bar{q} \leq q^t \\ y^{-1}(z(\bar{q})) + \bar{q} & \text{if } q^t < \bar{q} \leq \tilde{q} \end{cases} \quad (24)$$

The first-order derivatives can be shown to be given by the following theorem.

**Theorem 5:** Assuming fixed  $\{\eta^C, \eta^E\}$ , the partial derivatives of the surpluses  $S_M^E(\bar{q})$ ,  $S_M^C(\bar{q})$ ,  $S_M^I(\bar{q})$ ,  $S_M^T(\bar{q})$  with respect to  $\bar{q}$ , for  $0 \leq \bar{q} \leq \tilde{q}$ , is given by:

$$\begin{aligned} \frac{\partial S_M^E(\bar{q})}{\partial \bar{q}} &= \begin{cases} \frac{y\eta^E}{\eta^E - 1} & \text{if } \bar{q} \leq q^t \\ -\frac{p(\bar{q})y\eta^C}{q\eta^E} & \text{if } q^t < \bar{q} \leq \tilde{q} \end{cases} \\ \frac{\partial S_M^C(\bar{q})}{\partial \bar{q}} &= \begin{cases} \frac{y}{\mu - \bar{q}}[(v'(y) - \bar{q})\eta^E - (\mu - \bar{q})] & \text{if } \bar{q} \leq q^t \\ -y & \text{else} \end{cases} \\ \frac{\partial S_M^I(\bar{q})}{\partial \bar{q}} &= \begin{cases} y & \text{if } \bar{q} \leq q^t \\ \frac{y\eta^C}{\bar{q}}[\mu - \bar{q}(1 - 1/\eta^C) - p(\bar{q})(1 - 1/\eta^E)] & \text{else} \end{cases} \\ \frac{\partial S_M^T(\bar{q})}{\partial \bar{q}} &= \begin{cases} \frac{y\eta^E}{(\mu - \bar{q})} (v'(y) + \frac{\mu - \eta^E \bar{q}}{\eta^E - 1}) & \text{if } \bar{q} \leq q^t \\ -\frac{y\eta^C}{\bar{q}} (p(\bar{q}) + \bar{q} - \mu) & \text{if } q^t < \bar{q} \leq \tilde{q} \end{cases} \end{aligned}$$

where  $p(\bar{q}) = r(\bar{q}) - \bar{q}$  is the end-user price under content-provider price restriction  $\bar{q}$ , and  $y = y(p(\bar{q}))$  is the end-user demand at price  $p(\bar{q})$ .

It is clear that the ISP surplus increases with  $\bar{q}$ . For  $\bar{q} \leq q^t$ , the content-provider surplus increases provided the end-user price sensitivity is sufficiently high relative to the cost of provisioning capacity as in the competitive market case. For  $\bar{q} > q^t$ , the surplus decreases. The total surplus increases for  $\bar{q} < q^t$  and decreases for  $\bar{q} \geq q^t$ , as it can be shown that  $p(\bar{q}) + \bar{q} \geq \mu$ .

The end-user surplus increases for  $\bar{q} \leq q^t$  and decreases for  $\bar{q} > q^t$ . To quantify the surplus at  $\bar{q} = \tilde{q}$ , we re-write the monopoly price structure given by (15) as

$$\tilde{p} = \frac{\mu}{(1 - 1/\eta^E)} - \frac{\tilde{q}(1 - 1/\eta^C)}{(1 - 1/\eta^E)} \quad (25)$$

Notice that the first term on the right hand side is the monopoly price in the absence of content-provider participation. The presence of the content-provider reduces the end-user price and increases the end-user surplus, provided that the content-provider has an elasticity of demand greater than 1. If we allow content-provider elasticity of demand  $\eta^C < 1$ , then the end-user faces increased prices due to content-provider participation! When the content-provider is highly price insensitive, with an elasticity of demand  $\eta^C < 1$ , the monopoly ISP reaps all the

benefits from the content-provider's participation, and in fact extracts more from the end-user.

The monopoly ISP stands to gain in surplus from content-provider participation at all price restrictions  $0 \leq \bar{q} \leq \tilde{q}$ . What about the monopoly price? The equilibrium monopoly price is given by

$$\tilde{r} = \frac{\mu}{(1 - 1/\eta^E)} - \frac{\tilde{q}(1/\eta^E - 1/\eta^C)}{(1 - 1/\eta^E)} \quad (26)$$

where the first term is the monopoly price in the absence of the content-provider. It follows that a content-provider less price sensitive than the end-user ( $\eta^C < \eta^E$ ) increases the monopoly equilibrium price relative to the case where the content-provider does not participate. A more price sensitive content-provider, on the other hand, decreases the monopoly ISP equilibrium price.

Figure 4 is a plot of the end-user and content-provider surplus in a monopoly market against  $\bar{q}$  for the example illustrated in Section V with  $\{a_f = \$50, \alpha_f = 0.5, b_f = \$50, \beta_f = 0.5\}$ . Figure 5 plots the ISP and the total surplus for the same example.

In summary, in a monopoly ISP market:

- Content provider price restriction below a threshold price  $q^t \leq \bar{q}$  results in a drop in the end-user demand below that of content-provider demand.
- End user gains in surplus from easing price restriction for  $\bar{q} \leq q^t$  and loses in surplus for  $\bar{q} > q^t$ . The end-user surplus is maximized at  $\bar{q} = q^t$ . With no price restriction, end-user benefits from content-provider participation if the content-provider price sensitivity is greater than 1.
- The total surplus increases for  $\bar{q} \leq q^t$  and decreases for  $\bar{q} > q^t$ . The total surplus is maximized at  $\bar{q} = q^t$ .
- The content-provider surplus increases at low values of  $\bar{q}$  if the end-user price sensitivity is sufficiently high relative to the cost of connectivity. Lowering the cost of connectivity reduces the end-user price sensitivity required for the content-provider to gain from participation in rate allocation.
- The content-provider surplus decreases at high values of  $\bar{q}$ . The value of  $\bar{q}$  at which the content-provider surplus is maximized is typically different from  $q^t$ .
- The monopoly ISP price level increases with content-provider participation if the content-provider is less price sensitive than the end-user and decreases otherwise. Regardless of the change in the monopoly ISP price, the monopoly ISP surplus always increases with content-provider participation.

## VII. CONCLUSION

We presented a framework for network data rate allocation by considering the three-way interaction between end-users, ISPs, and content-providers. It yields quantitative results providing a precise and relevant “language” to structure one of the aspects in the on-going debate about “Network Neutrality”: rate allocation and surplus distribution with content provider pricing. The results from the developed framework indicate that the Internet connectivity market can benefit from content-provider participation but suffers from monopolistic ISP market conditions. In addition to the overall connectivity market,

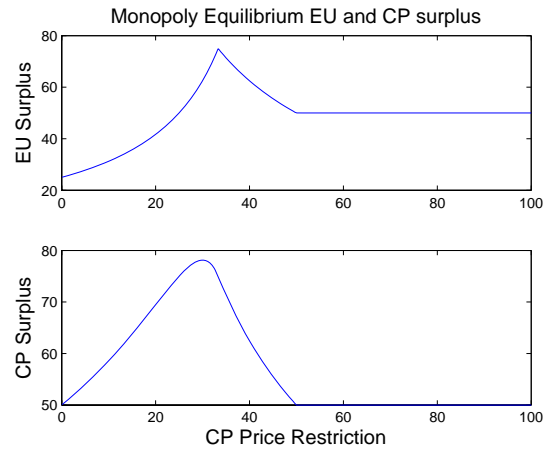


Fig. 4. Monopoly Equilibrium EU and CP Surplus vs.  $\bar{q}$ . End user surplus is maximized at  $\bar{q} = q^t$ , content-provider surplus is maximized at  $\bar{q}$  less than  $q^t$ .

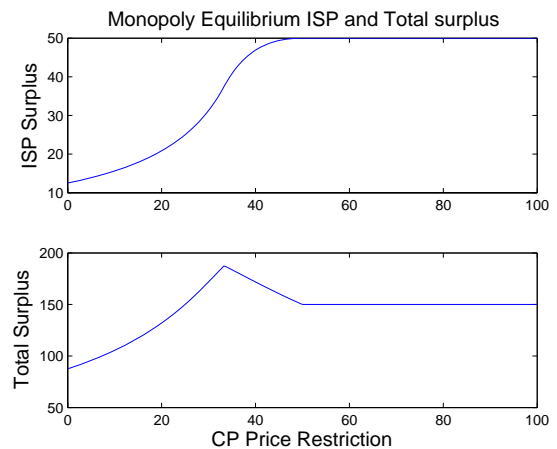


Fig. 5. Monopoly Equilibrium ISP and Total Surplus vs.  $\bar{q}$ . ISP surplus increases monotonically. Total surplus peaks at  $\bar{q} = q^t$ .

content-providers themselves can benefit from participation, if the cost of connectivity is low and the end-users are sufficiently price sensitive. Decreasing the cost of connectivity and increasing competitiveness in the ISP market, rather than restricting content-provider pricing a priori, appears to be an appropriate set of measures.

While we have presented a model that provides a useful tool to capture the complex interactions involved in the Internet connectivity market, further work is required to explicitly incorporate interactions that were abstracted out of this model. For example, a more complete picture will emerge if the interaction involving content distribution networks and peer-to-peer networks, as well as the transit and peering arrangements between ISPs [16] are incorporated into the model.

## APPENDIX A: PROOF OF THEOREM 1

Since  $y_f(p_f), z_f(q_f)$  are monotonically decreasing, the price structure  $\{\hat{p}_f, \hat{q}_f\}$  that maximizes (8) should satisfy  $\hat{p}_f + \hat{q}_f = \mu_f$ . Now it is necessarily the case that  $x_f(\mu_f) = y_f(\hat{p}_f) = z_f(\hat{q}_f)$ . If not, let  $y_f(\hat{p}_f) < z_f(\hat{q}_f)$ . Then  $\hat{p}_f > 0$  (if not, Assumption 2 is violated). This implies that  $\hat{p}_f$  can be decreased by a small amount and correspondingly  $\hat{q}_f$  increased



by a small amount resulting in a larger objective value for the maximization problem, contradicting that  $\{\hat{p}_f, \hat{q}_f\}$  is the optimum. The argument is similar if we let  $y_f(\hat{p}_f) > z_f(\hat{q}_f)$ . Further, we have  $u'_f(x_f) = \hat{p}_f$  and  $v'_f(x_f) = \hat{q}_f$  which implies that  $u'_f(x_f) + v'_f(x_f) = \mu_f$ , the condition required for maximization of system surplus.

## APPENDIX B: PROOF OF THEOREM 2

We suppress subscript  $f$  for notational simplicity. Let  $g(\theta) = y(\theta r) - z((1-\theta)r)$ . We assume bounded first derivatives for the demand functions. Since  $g(\theta)$  is monotonic decreasing and continuous in  $\theta \in [0, 1]$ , with  $g(0) > 0 > g(1)$ , there exists a unique  $\theta^* \in [0, 1]$  such that  $g(\theta^*) = 0$ .

Consider  $T$  such that  $|g'(\theta)| < T$ . For  $t > T$ , we have one of  $\theta[t] = \theta^*$ ,  $\theta[t] < \theta^*$ , or  $\theta[t] > \theta^*$ . First suppose that  $\theta[t] < \theta^*$ . Let  $\hat{\theta}[t+1] = \theta[t] + (1/t)g(\theta[t])$ , so that  $\theta[t+1] = B(\hat{\theta}[t+1])$ . Since  $g$  is monotonically decreasing,  $g(\theta[t]) > 0$ , so that  $\hat{\theta}[t+1] > \theta[t]$ . Further,  $\hat{\theta}[t+1] < \theta^*$ . To see this, note that

$$\begin{aligned} g(\hat{\theta}[t+1]) &\stackrel{(a)}{=} g(\theta[t]) + \frac{g'(\theta')}{t}g(\theta[t]) \\ &\stackrel{(b)}{\geq} g(\theta[t]) - \frac{|g'(\theta')|}{T}g(\theta[t]) \\ &\stackrel{(c)}{>} g(\theta[t]) - g(\theta[t]) = 0. \end{aligned} \quad (27)$$

where (a) follows from Taylor's theorem for some  $\theta'$ , (b) follows from the fact that  $t > T$  and (c) follows from the fact that  $|g'(\theta)| < T$  for all  $\theta$ . Since  $g$  is monotonically decreasing, we have that  $\hat{\theta}[t+1] < \theta^*$ . We have,  $\hat{\theta}[t+1] \in [\theta[t], \theta^*] \subset [0, 1]$ , and therefore,  $\theta[t+1] = B(\hat{\theta}[t+1]) = \hat{\theta}[t+1]$ . We have shown that if  $t > T$  and  $\theta[t] < \theta^*$ ,  $\theta[t+1] \in [\theta[t], \theta^*]$ . Therefore, for  $t > T$ ,  $\theta[t]$  is an increasing sequence bounded above by  $\theta^*$ , and hence converges to some  $\theta_\infty \leq \theta^*$ . Similar argument applies for  $\theta[t] > \theta^*$ , resulting in convergence to  $\theta_\infty \geq \theta^*$ .

It has to be shown that  $\theta_\infty = \theta^*$ . Suppose not, let  $\theta[T] < \theta_\infty < \theta^*$ . We have  $g(\theta_\infty) > g(\theta^*) = 0$ , and thus

$$\begin{aligned} \theta_\infty &= \theta[T] + \sum_{t=T}^{\infty} 1/tg(\theta[t]) \\ &> \theta[T] + g(\theta_\infty) \sum_{t=T}^{\infty} 1/t = \infty \end{aligned} \quad (28)$$

which is a contradiction. Similar argument applies to the case when  $\theta[T] > \theta_\infty > \theta^*$ .

## APPENDIX C: PROOF OF THEOREM 3

If  $\{\hat{p}_f, \hat{q}_f\}$  represent a price split at which the maximum for problem (8) is achieved, then it can be shown [20] that there exist Lagrangian dual variables  $\{\hat{\nu}_f, \hat{\lambda}_f\}$  such that

$$\left. \begin{aligned} y_f(\hat{p}_f) &= z_f(\hat{q}_f) \\ z'_f(\hat{q}_f)(1 - \hat{\nu}_f) &= -\hat{\lambda}_f = y'_f(\hat{p}_f)\hat{\nu}_f \end{aligned} \right\} \text{if } 1 > \hat{\nu}_f > 0$$

$$\left. \begin{aligned} y_f(\hat{p}_f) &\leq z_f(\hat{q}_f) \\ \hat{\lambda}_f &= 0 \end{aligned} \right\} \text{if } \hat{\nu}_f = 1$$

$$\left. \begin{aligned} y_f(\hat{p}_f) &\geq z_f(\hat{q}_f) \\ \hat{\lambda}_f &= 0 \end{aligned} \right\} \text{if } \hat{\nu}_f = 0$$

$$(29)$$

Assumption 2 implies that the only possible case is  $1 > \hat{\nu}_f > 0$  for which we have  $x_f(r_f) = y_f(p_f) = z_f(q_f)$  and  $z'_f(\hat{q}_f)(1 - \hat{\nu}_f) = -\hat{\lambda}_f = y'_f(\hat{p}_f)\hat{\nu}_f$ . From results in convex optimization [18], we have  $x'_f(r_f) = -\lambda_f$  so that

$$x'_f(r_f) = \frac{y'_f(p_f)z'_f(q_f)}{y'_f(p_f) + z'_f(q_f)} \quad (30)$$

Dividing both the sides by  $x_f(r_f)$  and multiplying by  $r_f$ , we have

$$\eta_f = \frac{\eta_f^E \eta_f^C r_f}{\eta_f^E q_f + \eta_f^C p_f} \quad (31)$$

Combining (14) and (31), we get the resulting equilibrium price structure in (15).

## ACKNOWLEDGMENTS

The authors would like to acknowledge helpful discussions with Vinayak Tripathi at Princeton University and George Kesidis at Penn State.

## REFERENCES

- [1] J. M. Peha, "The Benefits and Risks of Mandating Network Neutrality, and the Quest for a Balanced Policy", *34th Telecommunications Policy Research Conference*, Sept. 2006.
- [2] C. Yoo, "Network Neutrality and the Economics of Congestion", *Georgetown Law Journal*, Vol. 94, June 2006
- [3] A. Odlyzko, "Network Neutrality, Search Neutrality, and the Never-Ending Conflict between Efficiency and Fairness in Markets," January 27, 2008. Available at SSRN: <http://ssrn.com/abstract=1095350>
- [4] T. Wu, "The Broadband Debate: A User's Guide," *Journal of Telecommunications and High Technology Law*, 3(69): 69-96, 2004.
- [5] L. Waverman, "Two-Sided Telecom Markets and the Unintended Consequences of Business Strategy," *Competition Policy International*, Vol. 3, No. 1, Spring 2007. Available at SSRN: <http://ssrn.com/abstract=987895>
- [6] F. P. Kelly, "Charging and rate control for elastic traffic", *European Transactions on Telecommunications* Volume 8, pages 3337, Jan. 1997.
- [7] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability", *J. of Op. Res. Soc.*, vol. 49, no. 3, pp.237-252, March 1998.
- [8] N. Economides and J. Tag, "Net Neutrality on the Internet: A Two-sided Market Analysis," NET Institute Working Paper No. 07-45; NYU Law and Economics Research Paper 07-40. Available at SSRN: <http://ssrn.com/abstract=1019121>
- [9] J. Musacchio, G. Schwartz, and J. Walrand, "Network Neutrality and Provider Investment Incentives," *Asilomar Conference on Signals, Systems and Computers*, 2007. Nov. 2007
- [10] J. Rochet and J. Tirole, "Two-Sided Markets: A Progress Report," *RAND Journal of Economics*, vol. 37, no. 3, pp 645-667, 2006.
- [11] S. Shakkottai and R. Srikant, "Economics of Network Pricing with Multiple ISPs," in *IEEE/ACM Transactions on Networking*, Volume 14, Issue 6, Dec. 2006.
- [12] D. Acemoglu and A. Ozdaglar, "Competition and Efficiency in Congested Markets," *Mathematics of Operations Research*, vol. 32, no. 1, pp. 1-31, February 2007.
- [13] J. K. Mackie-Mason and H. Varian, "Pricing congestible network resources," *IEEE Journal on Selected Areas in Communication*, vol. 13, no. 7, pp. 1141- 1149, 1995.
- [14] G. Kesidis, A. Das and G. de Veciana, "On Flat-rate and Usage-based pricing for tiered commodity internet services", *CISS*, March 2008.
- [15] S. H. Low, "A Duality Model of TCP and Queue Management Algorithms", *IEEE/ACM Tran. Networking*, vol. 11, no. 4, pp. 525-536, August 2003.
- [16] M. Chiang, "Network distribution capacity and content-pipe gap", Proc. 42nd CISS, March 2008.
- [17] R. Srikant. *The Mathematics of Internet Congestion Control*. Birkhauser, 2004.
- [18] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [19] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic Theory*, Oxford University Press, Oxford, United Kingdom, 1995.
- [20] D. P. Bertsekas, *Nonlinear Programming*, 2nd Ed., Athena Scientific, 1999.